

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 18, 2004, 18:59:56 ; Search time 3974.06 Seconds
(without alignments)
16370.617 Million cell updates/sec

Title: US-09-845-416-10_COPY_1500_3000
Perfect score: 1501
Sequence: 1 acaatggcaagacctccaag.....accagcgagctggggcctc 1501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues 6940544

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_or.*
21: em_ov.*
22: em_ro.*
23: em_sts.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1501	100.0	5339	6	AX538620	Sequence
2	1501	100.0	5462	6	AX538621	Sequence
3	1501	100.0	5952	6	AR304538	Sequence
4	1501	100.0	5952	6	AX114289	Sequence
5	1501	100.0	8689	6	AX538622	Sequence
6	1501	100.0	11443	6	AX538624	Sequence
7	1501	100.0	12057	6	AX538627	Sequence
8	1501	100.0	12446	9	HSDMDR	Human mRNA
9	1501	100.0	13957	6	AX409637	Sequence
10	1501	100.0	13957	6	AX538581	Sequence
11	1501	100.0	13957	9	HUMDYS	MI8533 Homo sapien
12	1497.4	99.8	13977	4	AR220819	Sequence
13	1336.2	89.0	13887	4	AF070485	AF070485 Canis fam
14	1298.4	86.5	13815	6	AX306153	Sequence
15	1298.4	86.5	13815	6	AX538582	Sequence
16	1298.4	86.5	13815	10	MUSDYSA	Sequence
17	1298.4	86.5	13907	6	AR093392	Sequence
18	1298.4	86.5	13907	6	AR142592	Sequence
19	1044.6	69.6	13575	5	GGDYS	Me8859 Mouse dyster
20	985.2	65.6	4402	6	E30219	Sequence
21	916.6	61.1	5417	6	AX538619	Sequence
22	786	52.4	4402	6	E30220	Shortened d
23	704.6	46.9	2654	5	FSCDYSTRO	M37645 Torpedo cal
24	652	43.4	3521	5	AF339031	AF339031 Danio rer
25	650.4	43.3	4075	6	E30221	Shortened d
26	646.2	43.1	3747	6	E30218	Shortened d
27	619.2	41.3	10705	10	RNAJ2967	Sequence
28	619.2	41.3	11096	6	AX538584	Sequence
29	619.2	41.3	11096	6	AX538584	Sequence
30	617.8	41.2	3161	10	MMY12229	Sequence
31	604.8	40.3	6045	6	A63605	Sequence 7
32	604.8	40.3	6045	6	AR281528	Sequence
33	604.8	40.3	6059	6	AX107972	Sequence
34	604.8	40.3	10302	6	AX538583	Sequence
35	604.8	40.3	10302	9	HSMURS	X69086 H.sapiens m
36	604.8	40.3	10320	6	A63607	Sequence 9
37	604.8	40.3	10320	6	AR281529	Sequence
38	592.6	39.5	3163	6	E30223	Shortened d
39	534	35.6	3499	9	HGU43519	U43519 Human dyster
40	534	35.6	5106	6	AX552248	Sequence
41	533.2	35.5	3073	10	AF195788	AF195788 Rattus no
42	533.2	35.5	3140	10	AF195787	AF195787 Rattus no
43	500.6	33.4	4416	10	BC060605	BC060605 Mus muscu
44	484	32.2	2005	6	AX817318	AX817318 Sequence
45	484	32.2	2110	9	HUMDMDX	M92650 Human Duch

ALIGNMENTS

RESULT 1
AX538620

LOCUS AX538620
DEFINITION Sequence 40 from Patent WO0229056.
ACCESSION AX538620
VERSION AX538620.1 GI:25271166
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.

REFERENCE
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 40 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

AX538620 5339 bp DNA linear PAT 23-NOV-2002

FEATURES		Location/Qualifiers	
source		1..5339	
		/organism="synthetic construct"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:32630"	
		/note="Synthetic"	
ORIGIN			
Query Match		100.0%; Score 1501; DB 6; Length 5339;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 1501; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	ACAATGGCAGACCTCCAAAGTGAAATGAACTCACAAGATGTTTATCAACACCTGGA	60
Db	1693	ACAATGGCAGACCTCCAAAGTGAAATGAACTCACAAGATGTTTATCAACACCTGGA	1752
Qy	61	TGAAAAAGCCCAAAATTCCTGAGATCCCTGGAGGTTCCGATGATGCGAGTCCCTGTTACA	120
Db	1753	TGAAAAAGCCCAAAATTCCTGAGATCCCTGGAGGTTCCGATGATGCGAGTCCCTGTTACA	1812
Qy	121	AAGAGTTTGGATGAACATGAACTTCAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACAT	180
Db	1813	AAGAGTTTGGATGAACATGAACTTCAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACAT	1872
Qy	181	TAGTCCCATTTGGAAGCCAGTCTGACACCAAGTGAAGCGTCTGCACCTTTCTCTCGAGGA	240
Db	1873	TAGTCCCATTTGGAAGCCAGTCTGACACCAAGTGAAGCGTCTGCACCTTTCTCTCGAGGA	1932
Qy	241	ACTTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCCGCAGCGACCTATTGAGAG	300
Db	1933	ACTTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCCGCAGCGACCTATTGAGAG	1992
Qy	301	CGACTTTCAGCAGTTCAGAGCAGACAGATGTACATAGGCGCTTCAAGAGGGAATTGAA	360
Db	1993	CGACTTTCAGCAGTTCAGAGCAGACAGATGTACATAGGCGCTTCAAGAGGGAATTGAA	2052
Qy	361	AACATAAGAACCTGTAACTAGTACTCTTTGAGACTGTACGAATATTCTGCACAGCA	420
Db	2053	AACATAAGAACCTGTAACTAGTACTCTTTGAGACTGTACGAATATTCTGCACAGCA	2112
Qy	421	GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTCCCTCTGAGGAGAG	480
Db	2113	GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTCCCTCTGAGGAGAG	2172
Qy	481	AGCCAGAACTGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATACCTGAGTGGA	540
Db	2173	AGCCAGAACTGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATACCTGAGTGGA	2232
Qy	541	AAAATTGAACCTGACTCCCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT	600
Db	2233	AAAATTGAACCTGACTCCCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT	2292
Qy	601	CCAGGAATCTCAAGAGCCACGATGAGCTGGAACCTCAAGCTGGCCCAAGCTGAGGTGAT	660
Db	2293	CCAGGAATCTCAAGAGCCACGATGAGCTGGAACCTCAAGCTGGCCCAAGCTGAGGTGAT	2352
Qy	661	CAAGGATCTGCGAGCCCGTGGCGACTCTCTCATTTGACTCTCTCCAAAGATCACTCGA	720
Db	2353	CAAGGATCTGCGAGCCCGTGGCGACTCTCTCATTTGACTCTCTCCAAAGATCACTCGA	2412
Qy	721	GAAAGTCAAGGCACTTCGAGGAGAAATGCGCTCTGAAAGAGACGTCAGGACGTCGA	780
Db	2413	GAAAGTCAAGGCACTTCGAGGAGAAATGCGCTCTGAAAGAGACGTCAGGACGTCGA	2472
Qy	781	TGACCTTGTCTGCGACCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAAACCCTCAGC	840
Db	2473	TGACCTTGTCTGCGACCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAAACCCTCAGC	2532
Qy	841	TCTGGAGACCTTAACACAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCAG	900
Db	2533	TCTGGAGACCTTAACACAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCAG	2592
Qy	901	GCAGCTGTCATGAAGCCCAACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC	960
Db	2593	GCAGCTGTCATGAAGCCCAACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC	2652
Qy	961	TGTCCAGGCTTCCCTGGGAGAGCCATCTCCGCAAAAGAGTGCCTTACTATATCAACCA	1020
Db	2653	TGTCCAGGCTTCCCTGGGAGAGCCATCTCCGCAAAAGAGTGCCTTACTATATCAACCA	2712
Qy	1021	CGAGACTCAAAACAACTTCTCTGGGACCATCCCAAAATGACAGAGCTTACCACTTTTAGC	1080
Db	2713	CGAGACTCAAAACAACTTCTCTGGGACCATCCCAAAATGACAGAGCTTACCACTTTTAGC	2772
Qy	1081	TGACCTGAATATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAAGACTGCA	1140
Db	2773	TGACCTGAATATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAAGACTGCA	2832
Qy	1141	GAAAGCCCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGCATGTGATGCTTGGACCAAGCA	1200
Db	2833	GAAAGCCCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGCATGTGATGCTTGGACCAAGCA	2892
Qy	1201	CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGACATTAATTTGTTTGAACAC	1260
Db	2893	CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGACATTAATTTGTTTGAACAC	2952
Qy	1261	TATTTATGACCCCTGGAGCAAGAGCACAAATTTGTTTGAACATTAATTTGTTTGAACAC	1320
Db	2953	TATTTATGACCCCTGGAGCAAGAGCACAAATTTGTTTGAACATTAATTTGTTTGAACAC	3012
Qy	1321	TATGTGTCTGAACCTGGCTGTGAATTTTATGATACGGGACGAACAGGAGGATCCGTGT	1380
Db	3013	TATGTGTCTGAACCTGGCTGTGAATTTTATGATACGGGACGAACAGGAGGATCCGTGT	3072
Qy	1381	CCTCTCTTTTAAACTGGCATCTTCCCTGTGTAAAGACATTTGGAACACAGTACAG	1440
Db	3073	CCTCTCTTTTAAACTGGCATCTTCCCTGTGTAAAGACATTTGGAACACAGTACAG	3132
Qy	1441	ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACGAGCGCAGGCTGGGCT	1500
Db	3133	ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACGAGCGCAGGCTGGGCT	3192
Qy	1501	C 1501	
Db	3193	C 3193	
RESULT 2			
AX538621			
LOCUS			
AX538621			
DEFINITION			
Sequence 41 from Patent WO0229056.			
ACCESSION			
AX538621			
VERSION			
AX538621.1			
KEYWORDS			
synthetic construct			
SOURCE			
ORGANISM			
synthetic construct			
artificial sequences.			
REFERENCE			
1			
AUTHORS			
Chamberlain,J.S. and Harper,S.Q.			
TITLE			
Mini-dystrophin nucleic acid and peptide sequences			
JOURNAL			
Patent: WO 0229056-A 41 11-APR-2002;			
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)			
FEATURES			
source			
1..5462			
/organism="synthetic construct"			
/mol_type="unassigned DNA"			
/db_xref="taxon:32630"			
/note="Synthetic"			
ORIGIN			
Query Match			
100.0%; Score 1501; DB 6; Length 5462;			
Best Local Similarity			
100.0%; Pred. No. 0;			
Matches 1501; Conservative			
0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ACAATGCAAGACCTCCAAAGTGAATTTGAAGTCAACAGATGTTTATCAACACCTGGA	60

Db 3342 ACTTCTGGTGTGCTACAGCTGAAGATGATGAATTAAGCCGCGACGACCTATTGAGG 3401
QY 301 CGACTTTCCAGCACTTCAAGACGACGATGATACATAGAGCCCTTCAAGAGGAATTGAA 360
Db 3402 CGACTTTCCAGCACTTCAAGACGACGATGATACATAGAGCCCTTCAAGAGGAATTGAA 3461
QY 361 AACTAAAGAACTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGCA 420
Db 3462 AACTAAAGAACTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGCA 3521
QY 421 GCCTTTGGAAGGACTAGAGAACTCTACAGAGGCCAGAGACTGCTCTCTGAGAGAG 480
Db 3522 GCCTTTGGAAGGACTAGAGAACTCTACAGAGGCCAGAGACTGCTCTCTGAGAGAG 3581
QY 481 AGCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGG 540
Db 3582 AGCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGG 3641
QY 541 AAAATTGAACCTGCACTCCGCTCACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACT 600
Db 3642 AAAATTGAACCTGCACTCCGCTCACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACT 3701
QY 601 CCAGGAATTCAGAGCCACGATGAGTGGACCTCAAGCTGCGCAAGCTGAGTGTAT 560
Db 3702 CCAGGAATTCAGAGCCACGATGAGTGGACCTCAAGCTGCGCAAGCTGAGTGTAT 3761
QY 661 CAAGGGATCTGGCAGCCCTGGGCGATCTCTCTATTGACTCTCTCCAAAGATCACTCGA 720
Db 3762 CAAGGGATCTGGCAGCCCTGGGCGATCTCTCTATTGACTCTCTCCAAAGATCACTCGA 3821
QY 721 GAAAGTCAAGGCACTTCAGAGCAAAATGGCCCTCTGAAAGAGAGCTGAGCCAGCTCAA 780
Db 3822 GAAAGTCAAGGCACTTCAGAGCAAAATGGCCCTCTGAAAGAGAGCTGAGCCAGCTCAA 3881
QY 781 TGACCTTGCTCGCAGCTTACCACTTTGGCATTTCAGCTCTCACCTATAAATCACTCAGAC 840
Db 3882 TGACCTTGCTCGCAGCTTACCACTTTGGCATTTCAGCTCTCACCTATAAATCACTCAGAC 3941
QY 841 TCTGGAAGACCTGAACACAGATGGAAGCTTTCAGAGTGGCCGCTGAGGACCGAGTCA 900
Db 3942 TCTGGAAGACCTGAACACAGATGGAAGCTTTCAGAGTGGCCGCTGAGGACCGAGTCA 4001
QY 901 GCAGCTGCATGAAGCCACAGGACTTTCGTCAGCATCTCAGCATCTTCTTCCACGTC 960
Db 4002 GCAGCTGCATGAAGCCACAGGACTTTCGTCAGCATCTCAGCATCTTCTTCCACGTC 4061
QY 961 TGTCCAGGCTTCTGGGAGAGACCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 1020
Db 4062 TGTCCAGGCTTCTGGGAGAGACCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 4121
QY 1021 CGAGACTCAAAACAACTTCTGGGACCATCCCAAAATGACAGACTTACCAGTCTTTAGC 1080
Db 4122 CGAGACTCAAAACAACTTCTGGGACCATCCCAAAATGACAGACTTACCAGTCTTTAGC 4181
QY 1081 TGACCTGAATATGTCAGATTTCTCAGCTTATAGGACTGCGCATGAAACTCCGAAAGACTGCA 1140
Db 4182 TGACCTGAATATGTCAGATTTCTCAGCTTATAGGACTGCGCATGAAACTCCGAAAGACTGCA 4241
QY 1141 GAAGGCCCTTGTGTTGATCTCTGAGCCTGTCAGCTGATGATGCTTGGACCA 1200
Db 4242 GAAGGCCCTTGTGTTGATCTCTGAGCCTGTCAGCTGATGATGCTTGGACCA 4301
QY 1201 CAACCTCAAGCAAAATGACAGCCATGATATCTCGAGATTTATTAATGTTTGACCA 1260
Db 4302 CAACCTCAAGCAAAATGACAGCCATGATATCTCGAGATTTATTAATGTTTGACCA 4361
QY 1261 TATTTATGACCGCTGGAGCAAGAGCAACAAATTTGTCACGCTCCCTCTCTCGGTGA 1320
Db 4362 TATTTATGACCGCTGGAGCAAGAGCAACAAATTTGTCACGCTCCCTCTCTCGGTGA 4421
QY 1321 TATGTCGTGAATGCTGCTCAATGTTTATGATACGGGACGAAACAGGGAGGATCCGCT 1380
Db 4422 TATGTCGTGAATGCTGCTCAATGTTTATGATACGGGACGAAACAGGGAGGATCCGCT 4481

QY 1381 CTTGCTTTTAAACTGGCATTTTCCCTGTGTAAGACCATTTTGAAGACAAAGTACAG 1440
Db 4482 CTTGCTTTTAAACTGGCATTTTCCCTGTGTAAGACCATTTTGAAGACAAAGTACAG 4541
QY 1441 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACAGCGCAGGCTGGGCT 1500
Db 4542 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACAGCGCAGGCTGGGCT 4601
QY 1501 C 1501
Db 4602 C 4602

RESULT 4
AX114289
LOCUS AX114289 5952 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 1 from Patent WO0129243.
ACCESSION AX114289
VERSION AX114289.1 GI:14031259
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Paul, X.L. and Xiao, X.
Method and vector for producing and transferring trans-spliced peptides
Patent: WO 0129243-A 1 26-APR-2001;
DALHOUSIE UNIVERSITY (CA); UNIV. OF PITTSBURGH OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION (US)
FEATURES
source 1. .5952
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
misc_feature 2897..2898
/note="S4 junction site"
misc_feature 3198..3199
/note="S2 junction site"

ORIGIN
Query Match 100.0%; Score 1501; DB 6; Length 5952;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAATGCAAGACTTCAAGGTGAAATTTGAAGCTCACAGATGTTTATCACAACTGGA 60
Db 3102 ACAAATGCAAGACTTCAAGGTGAAATTTGAAGCTCACAGATGTTTATCACAACTGGA 3161
QY 61 TGAACACAGCAAAATTTCTGAGATCCCTGGAAGTTCGGATGATGAGTCTGTTTACA 120
Db 3162 TGAACACAGCAAAATTTCTGAGATCCCTGGAAGTTCGGATGATGAGTCTGTTTACA 3221
QY 121 AAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAAGTTCGGAAGAAAGTCTTCAACAT 180
Db 3222 AAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAAGTTCGGAAGAAAGTCTTCAACAT 3281
QY 181 TAGGTCCTTGGATAAGCCAGTTCGACCAAGTGAAGCGTCTGCACTTTCTCTGAGCA 240
Db 3282 TAGGTCCTTGGATAAGCCAGTTCGACCAAGTGAAGCGTCTGCACTTTCTCTGAGCA 3341
QY 241 ACTTCTGCTGCTGCTACAGCTGAAAGATGATGAATTAACCGCGAGGACCTATTGGAGG 300
Db 3342 ACTTCTGCTGCTGCTACAGCTGAAAGATGATGAATTAACCGCGAGGACCTATTGGAGG 3401
QY 301 CGACTTTCCAGCACTTCAAGAGCAAGACGATGATACATAGGGCTTCAAGAGGAAATTGAA 360
Db 3402 CGACTTTCCAGCACTTCAAGAGCAAGACGATGATACATAGGGCTTCAAGAGGAAATTGAA 3461
QY 361 AACTAAAGAACTGTAATCATGAGTACTCTTGGAGTGTACGAATATTTCTGACAGCA 420

Db	3462	AACTAAAGAACTCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA	3521
Qy	421	GCCTTTGAAGGACTAGAGAACTCTACAGAGCCAGAGAGTGCCTCTGAGGAGAG	480
Db	3522	GCCTTTGAAGGACTAGAGAACTCTACAGAGCCAGAGAGTGCCTCTGAGGAGAG	3581
Qy	481	AGCCACAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA	540
Db	3582	AGCCACAGAAATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA	3641
Qy	541	AAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAATAATAGATGAGACCCCTTGAAGACT	600
Db	3642	AAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAATAATAGATGAGACCCCTTGAAGACT	3701
Qy	601	CCAGGAATCTCAAGAGGCCACGGATGAGCTGACCTCAAGCTGCGCCCAAGCTGAGGTGAT	660
Db	3702	CCAGGAATCTCAAGAGGCCACGGATGAGCTGACCTCAAGCTGCGCCCAAGCTGAGGTGAT	3761
Qy	661	CAAGGGATCTCGCAGCGCGTGGGGATCTCTCTAATGACTCTCTCCAAAGTACACCTGGA	720
Db	3762	CAAGGGATCTCGCAGCGCGTGGGGATCTCTCTAATGACTCTCTCCAAAGTACACCTGGA	3821
Qy	721	GAAAGTCAAGGACACTTCGAGGAGAAATTCGGCTCTGAAAGAGACGTGAGCCAGTCAA	780
Db	3822	GAAAGTCAAGGACACTTCGAGGAGAAATTCGGCTCTGAAAGAGACGTGAGCCAGTCAA	3881
Qy	781	TGACCTTGTCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC	840
Db	3882	TGACCTTGTCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC	3941
Qy	841	TCGGAAGACCTGAAACACAGATGGAAGCTTCTGAGTGGCGCTGAGGACCGAGTCAAG	900
Db	3942	TCGGAAGACCTGAAACACAGATGGAAGCTTCTGAGTGGCGCTGAGGACCGAGTCAAG	4001
Qy	901	CGAGCTGCATGAAGCCCAAGGACCTTTGGTCTCAGCATCTCAGCACTTCTTTTCCACGTC	960
Db	4002	CGAGCTGCATGAAGCCCAAGGACCTTTGGTCTCAGCACTTCTTTTCCACGTC	4061
Qy	961	TGTCCAGGTCCTGCGGAGAGGCCATCTCGCCAAACAAAGTGCCTCTACTATCAAGCA	1020
Db	4062	TGTCCAGGTCCTGCGGAGAGGCCATCTCGCCAAACAAAGTGCCTCTACTATCAAGCA	4121
Qy	1021	CGAGCTCAAAACAACTTGTGGAGCCATCCCAAAATGACAGAGCTCTACAGCTCTTTAGC	1080
Db	4122	CGAGCTCAAAACAACTTGTGGAGCCATCCCAAAATGACAGAGCTCTACAGCTCTTTAGC	4181
Qy	1081	TGACCTGAATAATGTGAGATTTCTCAGCTTATAGGACTGCCATGAACTCCGAAAGCTGCA	1140
Db	4182	TGACCTGAATAATGTGAGATTTCTCAGCTTATAGGACTGCCATGAACTCCGAAAGCTGCA	4241
Qy	1141	GAAGGCCCTTTGCTTGTGATCTTTGAGCCCTGACGTCATGTGATGCCCTTGGACGAGCA	1200
Db	4242	GAAGGCCCTTTGCTTGTGATCTTTGAGCCCTGACGTCATGTGATGCCCTTGGACGAGCA	4301
Qy	1201	CAACCTCAAGCAAAATGACAGCCCATGGATATCTTGCAGATTAATTAATTTGTTGACCA	1260
Db	4302	CAACCTCAAGCAAAATGACAGCCCATGGATATCTTGCAGATTAATTAATTTGTTGACCA	4361
Qy	1261	TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAAGCTCCCTCTCTGCGTGA	1320
Db	4362	TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAAGCTCCCTCTCTGCGTGA	4421
Qy	1321	TATGTCGTAAGCTGGCTGCTGAATGTTTATGATACGGACGAAACAGGAGGATCCGTGT	1380
Db	4422	TATGTCGTAAGCTGGCTGCTGAATGTTTATGATACGGACGAAACAGGAGGATCCGTGT	4481
Qy	1381	CTGTCTCTTTTAAACTGGCATATTTCCCTGTGTAAAGCAATTTGGAAGACAAAGTACAG	1440
Db	4482	CTGTCTCTTTTAAACTGGCATATTTCCCTGTGTAAAGCAATTTGGAAGACAAAGTACAG	4541
Qy	1441	ATACCTTTTCAAGCAAGTGCAGTCTCAACAGATTTTGTGACAGCGAGGCTGGGCT	1500
Db	4542	ATACCTTTTCAAGCAAGTGCAGTCTCAACAGATTTTGTGACAGCGAGGCTGGGCT	4601

Qy

1501

C

1501

Db

4602

C

4602

RESULT 5

AX538622

LOCUS

AX538622

DEFINITION

Sequence 42 from Patent WO0229056.

ACCESSION

AX538622

VERSION

AX538622.1

GI:25271171

KEYWORDS

synthetic construct

SOURCE

synthetic construct

ORGANISM

artificial sequences.

REFERENCE

1

AUTHORS

Chamberlain, J.S. and Harper, S.O.

TITLE

Mini-dystrophin nucleic acid and peptide sequences

JOURNAL

Patent: WO 0229056-A 42 11-APR-2002;

THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

FEATURES

Location/Qualifiers

1..8689

source

1..8689

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="Synthetic"

ORIGIN

Query Match

100.0%;

Score 1501;

DB 6;

Length 8689;

Best Local Similarity

100.0%;

Pred. No. 0;

Matches 1501;

Conservative 0;

Mismatches 0;

Indels 0;

Gaps 0;

Qy

1

ACAATGGCAAGACCTCCAAAGTGAATTAAGCTCACAAGATGTTTATCACACCTGGA

60

Db

3148

ACAATGGCAAGACCTCCAAAGTGAATTAAGCTCACAAGATGTTTATCACACCTGGA

3207

Qy

61

TGAAAAACAGCAAAAAATCTGAGATCCCTGGAAGGTTCGATGATCGAGTCTCTGTTCAC

120

Db

3208

TGAAAAACAGCAAAAAATCTGAGATCCCTGGAAGGTTCGATGATCGAGTCTCTGTTCAC

3267

Qy

121

AAGACGTTTGGATTAACATGAACATTCAGTGGAGTGAATCTCGGAAAAAGTCTCTCAACAT

180

Db

3268

AAGACGTTTGGATTAACATGAACATTCAGTGGAGTGAATCTCGGAAAAAGTCTCTCAACAT

3327

Qy

181

TAGTCCCATTTGAAACCCAGTCTGACCATGGAAGCTGCAACCTTTCTCTGCAAGGA

240

Db

3328

TAGTCCCATTTGAAACCCAGTCTGACCATGGAAGCTGCAACCTTTCTCTGCAAGGA

3387

Qy

241

ACTTCTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGGCACTTATTTGAGG

300

Db

3388

ACTTCTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGGCACTTATTTGAGG

3447

Qy

301

CGACTTTCCAGAGTTCAGAGCAAGCATGTACATAGGCGCTTCAAGAGGGAATTGAA

360

Db

3448

CGACTTTCCAGAGTTCAGAGCAAGCATGTACATAGGCGCTTCAAGAGGGAATTGAA

3507

Qy

361

AACTAAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA

420

Db

3508

AACTAAAGAACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA

3567

Qy

421

GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCTCTGAGGAGAG

480

Db

3568

GCCTTTGGAAGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTGCTCTGAGGAGAG

3627

Qy

481

AGCCCAAGATGCTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA

540

Db

3628

AGCCCAAGATGCTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA

3687

Qy

541

AAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT

600

Db

3688

AAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT

3747

```
QY 601 CCAGGAATCTCAAGAGCCACGGATGAGTGGACCTCAAGCTGGCCAGCAAGCTGAGGTGAT 660
Db 3748 CCAGGAATCTCAAGAGCCACGGATGAGTGGACCTCAAGCTGGCCAGCAAGCTGAGGTGAT 3807
QY 661 CAAGGGATCTTGGAGCCCGTGGCGGATCTCTCAATGACTCTCTCAAGATCAACCTCGA 720
Db 3808 CAAGGGATCTTGGAGCCCGTGGCGGATCTCTCAATGACTCTCTCAAGATCAACCTCGA 3867
QY 721 GAAAGTCAAGGCACTTCGAGGGAATTCGGCTCTGAAAGAAAGCTGAGCCAGCTCAA 780
Db 3868 GAAAGTCAAGGCACTTCGAGGGAATTCGGCTCTGAAAGAAAGCTGAGCCAGCTCAA 3927
QY 781 TGACCTTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAAGCTCAGAC 840
Db 3928 TGACCTTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAAGCTCAGAC 3987
QY 841 TCTGGAGACCTGAACACACGATGGAAGCTTCTGAGGTTGGCGTTCGAGGACCGAGTCA 900
Db 3988 TCTGGAGACCTGAACACACGATGGAAGCTTCTGAGGTTGGCGTTCGAGGACCGAGTCA 4047
QY 901 GCAGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGCACTTCTCTTCCACGTC 960
Db 4048 GCAGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGCACTTCTCTTCCACGTC 4107
QY 961 TGTCCAGGTTCCCTGGGAGAGCCATCTGGCCAAACAAAGTGCCTACTATATCAACCA 1020
Db 4108 TGTCCAGGTTCCCTGGGAGAGCCATCTGGCCAAACAAAGTGCCTACTATATCAACCA 4167
QY 1021 CGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAC 1080
Db 4168 CGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAC 4227
QY 1081 TGACCTGAATATGTGATCTCAGCTTATAGGACTGCGATGAAACTCCGAGAGCTGCA 1140
Db 4228 TGACCTGAATATGTGATCTCAGCTTATAGGACTGCGATGAAACTCCGAGAGCTGCA 4287
QY 1141 GAAGGCCCTTGTGGATCTCTTGAGCCTGTGAGTGTGATGATGATGATGATGATGAT 1200
Db 4288 GAAGGCCCTTGTGGATCTCTTGAGCCTGTGAGTGTGATGATGATGATGATGATGAT 4347
QY 1201 CAACCTCAAGCAAAATGACAGCCATGATGATCTCTGAGATTAATTAATTTGTTGACCA 1260
Db 4348 CAACCTCAAGCAAAATGACAGCCATGATGATCTCTGAGATTAATTAATTTGTTGACCA 4407
QY 1261 TATTTATGACCGCTGGAGCAAGCAACAAATTTGTTGATCAACGCTCTCTCTGCGTGA 1320
Db 4408 TATTTATGACCGCTGGAGCAAGCAACAAATTTGTTGATCAACGCTCTCTCTGCGTGA 4467
QY 1321 TATGTCTGAACTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
Db 4468 TATGTCTGAACTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4527
QY 1381 CTGTCTTTTAAACTGGCATCATTTCCCTGTTGATGATGATGATGATGATGATGATGAT 1440
Db 4528 CTGTCTTTTAAACTGGCATCATTTCCCTGTTGATGATGATGATGATGATGATGATGAT 4587
QY 1441 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCTGGGCT 1500
Db 4588 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCTGGGCT 4647
QY 1501 C 1501
Db 4648 C 4648
```

RESULT 6
AX538624
LOCUS
DEFINITION
Sequence 44 from Patent WO0229056.
ACCESSION
AX538624
VERSION
AX538624.1 GI:25271175
KEYWORDS
synthetic construct

```
ORGANISM synthetic construct  
artificial sequences.  
REFERENCE 1  
AUTHORS Chamberlain,J.S. and Harper,S.Q.  
TITLE Mini-dysprophic nucleic acid and peptide sequences  
JOURNAL Patent: WO 0229056-A 44 11-APR-2002;  
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)  
FEATURES  
source location/Qualifiers  
1..11443  
/organism="synthetic construct"  
/mol_type="unassigned DNA"  
/db_xref="taxon:32630"  
/note="Synthetic"  
ORIGIN  
Query Match 100.0%; Score 1501; DB 6; Length 11443;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACATGCAAGACCTCCAGGTGAATTTGAAGCTCACAGATGTTTATCACACCTGGA 60  
Db 5902 ACATGCAAGACCTCCAGGTGAATTTGAAGCTCACAGATGTTTATCACACCTGGA 5961  
QY 61 TGAACACAGCAAAATTCCTGAGATCCCTTGAAGGTTCGGATGATCGAGTCTCTTACA 120  
Db 5962 TGAACACAGCAAAATTCCTGAGATCCCTTGAAGGTTCGGATGATCGAGTCTCTTACA 6021  
QY 121 AAGACGTTTCGATTAACATGAATTTCAAGTGGAGTGAATTCGGAAGAAAGCTCTCAACAT 180  
Db 6022 AAGACGTTTCGATTAACATGAATTTCAAGTGGAGTGAATTCGGAAGAAAGCTCTCTCAACAT 6081  
QY 181 TAGTCCCAATTTGGAAGCCAGTTCTGACCAAGTGGAAAGCTCTGACCTTTCTCTGAGGA 240  
Db 6082 TAGTCCCAATTTGGAAGCCAGTTCTGACCAAGTGGAAAGCTCTGACCTTTCTCTGAGGA 6141  
QY 241 ACTTCTGTTGCTGATCAGCTGAAGATGATGAATTAAGCCGCGAGGACCTATTGGAGG 300  
Db 6142 ACTTCTGTTGCTGATCAGCTGAAGATGATGAATTAAGCCGCGAGGACCTATTGGAGG 6201  
QY 301 CGACTTTCCAGCAGTTTCAGAGCAGAACCATGATGATAGGCTTCAAGAGGAAATTGA 360  
Db 6202 CGACTTTCCAGCAGTTTCAGAGCAGAACCATGATGATAGGCTTCAAGAGGAAATTGA 6261  
QY 361 AACTAAAGAACCTGTAAATCATGATGATCTTTGAGACTGATGAAATTTCTGACAGACA 420  
Db 6262 AACTAAAGAACCTGTAAATCATGATGATCTTTGAGACTGATGAAATTTCTGACAGACA 6321  
QY 421 GCCTTTGGAAGGACTAGAGAACTCTTACAGAGGACCGAGAGCTGCTCTCTGAGGAG 480  
Db 6322 GCCTTTGGAAGGACTAGAGAACTCTTACAGAGGACCGAGAGCTGCTCTCTGAGGAG 6381  
QY 481 AGCCCAAGATGTCACTCGCTTCTACGAAAGCAGGCTGAGGAGTCAATTAATCTGAGTGA 540  
Db 6382 AGCCCAAGATGTCACTCGCTTCTACGAAAGCAGGCTGAGGAGTCAATTAATCTGAGTGA 6441  
QY 541 AAAATTTGAACCTGCATCCGCTGATGAGCAGAGAAATAGATGAGACCTTTGAAGACT 600  
Db 6442 AAAATTTGAACCTGCATCCGCTGATGAGCAGAGAAATAGATGAGACCTTTGAAGACT 6501  
QY 601 CCAGGAATCTCAAGAGCCACGGATGAGTGGACCTCAAGCTGGCCAGCAAGCTGAGGTGAT 660  
Db 6502 CCAGGAATCTCAAGAGCCACGGATGAGTGGACCTCAAGCTGGCCAGCAAGCTGAGGTGAT 6561  
QY 661 CAAGGGATCTTGGAGCCCGTGGCGGATCTCTCAATGACTCTCTCAAGATCAACCTCGA 720  
Db 6562 CAAGGGATCTTGGAGCCCGTGGCGGATCTCTCAATGACTCTCTCAAGATCAACCTCGA 6621  
QY 721 GAAAGTCAAGGCACTTCGAGGGAATTCGGCTCTGAAAGAAAGCTGAGCCAGCTCAA 780  
Db 6622 GAAAGTCAAGGCACTTCGAGGGAATTCGGCTCTGAAAGAAAGCTGAGCCAGCTCAA 6681  
QY 781 TGACCTTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAAGCTCAGAC 840
```


Query Match		100.0%;	Score 1501;	DB 9;	Length 12446;	
Best Local Similarity		100.0%;	Pred. No. 0;			
Matches 1501;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	ACAATGGCAAGACCTTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATACAAACCTGGA	60			
Db	8306	ACAATGGCAAGACCTTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATACAAACCTGGA	8385			
QY	61	TGAAAAAGCAAAAAAATCTTGAGTCCCTGGAGGTTCCGATGATGAGTCTCTGTTACA	120			
Db	8366	TGAAAAAGCAAAAAAATCTTGAGTCCCTGGAGGTTCCGATGATGAGTCTCTGTTACA	8425			
QY	121	AAGACCTTTGGATAACATGAACCTTCAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACAT	180			
Db	8426	AAGACCTTTGGATAACATGAACCTTCAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACAT	8485			
QY	181	TAGTCCCATTTGGAAGCAGTTCTGACACAGTGGAGGCTTGACCTTCTCTGACGGA	240			
Db	8486	TAGTCCCATTTGGAAGCAGTTCTGACACAGTGGAGGCTTGACCTTCTCTGACGGA	8545			
QY	241	ACTTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG	300			
Db	8546	ACTTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG	8605			
QY	301	CGACTTTCCAGCAGTTCAGACGACGATGTATCATAGGGCTTCAAGAGGGAAATTGAA	360			
Db	8606	CGACTTTCCAGCAGTTCAGACGACGATGTATCATAGGGCTTCAAGAGGGAAATTGAA	8665			
QY	361	ACTAAAGACCTGTAAATCATGACTCTTGAGACTGTAGATATTTCTGACAGAGCA	420			
Db	8666	ACTAAAGACCTGTAAATCATGACTCTTGAGACTGTAGATATTTCTGACAGAGCA	8725			
QY	421	GCCTTTGGAAGGACTAGAAAACTCTACAGAGCCGACAGAGCTGCCTCTGAGGAGAG	480			
Db	8726	GCCTTTGGAAGGACTAGAAAACTCTACAGAGCCGACAGAGCTGCCTCTGAGGAGAG	8785			
QY	481	AGCCAGATGTCACTCGCGCTTACGAAAGCAGCTGAGGAGTCAATAGTGTGGA	540			
Db	8786	AGCCAGATGTCACTCGCGCTTACGAAAGCAGCTGAGGAGTCAATAGTGTGGA	8845			
QY	541	AAAAATTGAACCTGCATCCGCTGACTGGCAGAGAAAATAGATGAGACCTTGAAGACT	600			
Db	8846	AAAAATTGAACCTGCATCCGCTGACTGGCAGAGAAAATAGATGAGACCTTGAAGACT	8905			
QY	601	CCAGGAACCTCAAGAGCCACCGATGAGCTGACCTCAAGCTGCGCCCAAGCTGAGGTAT	660			
Db	8906	CCAGGAACCTCAAGAGCCACCGATGAGCTGACCTCAAGCTGCGCCCAAGCTGAGGTAT	8965			
QY	661	CAAGGATCTGCGACCCGCTGGCGATCTCTCTATTGACTCTCTCCAAAGATCACTCGA	720			
Db	8966	CAAGGATCTGCGACCCGCTGGCGATCTCTCTATTGACTCTCTCCAAAGATCACTCGA	9025			
QY	721	GAAAGTCAAGGCACTTTCGAGGAGAAATTCGGCTCTGAAAGAGAACGTGAGCCAGTCAA	780			
Db	9026	GAAAGTCAAGGCACTTTCGAGGAGAAATTCGGCTCTGAAAGAGAACGTGAGCCAGTCAA	9085			
QY	781	TGACCTTGTCTGCGACGTTACCACTTTGGCGATTCAGCTCTCACCGTATAAAGTCAAGC	840			
Db	9086	TGACCTTGTCTGCGACGTTACCACTTTGGCGATTCAGCTCTCACCGTATAAAGTCAAGC	9145			
QY	841	TCGGAAGACCTGAAACACAGATGGAAGTCTTCGAGGTGGCGTGGCGACCGAGTCAAG	900			
Db	9146	TCGGAAGACCTGAAACACAGATGGAAGTCTTCGAGGTGGCGTGGCGACCGAGTCAAG	9205			
QY	901	GCAGCTGCATGAAGCCACAGGACCTTTGGTCAGATCTCAGACCTTTCTTTCACGTC	960			
Db	9206	GCAGCTGCATGAAGCCACAGGACCTTTGGTCAGATCTCAGACCTTTCTTTCACGTC	9265			
QY	961	TGTCAGGGTCCCTGGAGAGCCATCTCGCAAAACAAAGTCCCTACTATATCAACCA	1020			
Db	9266	TGTCAGGGTCCCTGGAGAGCCATCTCGCAAAACAAAGTCCCTACTATATCAACCA	9325			
QY	1021	CGAGACTCAAAACAACTTGTCTGGGACCATCCAAAAATGACAGAGCTCTACAGTCTTTAGC	1080			

Db	9326	CGAGACTCAAAACAACCTTGTCTGGGACCATCCAAAAATGACAGAGCTCTACAGTCTTTAGC	9385			
QY	1081	TGACCTGATAATGTGACAGATTTCTCAGCTTATAGGACTGCATGAAACTCCGAGACTGCA	1140			
Db	9386	TGACCTGATAATGTGACAGATTTCTCAGCTTATAGGACTGCATGAAACTCCGAGACTGCA	9445			
QY	1141	GAAAGCCCTTTGCTTGGATCTCTTGGAGCCTGTGAGCTGCATGATGCTTGGACAGCA	1200			
Db	9446	GAAAGCCCTTTGCTTGGATCTCTTGGAGCCTGTGAGCTGCATGATGCTTGGACAGCA	9505			
QY	1201	CAACCTCAAGCAAAATGACAGCCCATGATATCTGAGATTTAAATGTTTGGACAC	1260			
Db	9506	CAACCTCAAGCAAAATGACAGCCCATGATATCTGAGATTTAAATGTTTGGACAC	9565			
QY	1261	TATTTATGACCGCTCGAGCAAGACCAACAATTTGGTCAACGTCCTCTCTGCGTGA	1320			
Db	9566	TATTTATGACCGCTCGAGCAAGACCAACAATTTGGTCAACGTCCTCTCTGCGTGA	9625			
QY	1321	TATGTGTGAACTGCTGCTGATTTTATGATACGGACGAAACAGGAGGATCCGTGT	1380			
Db	9626	TATGTGTGAACTGCTGCTGATTTTATGATACGGACGAAACAGGAGGATCCGTGT	9685			
QY	1381	CTGTCTTTTAAACTGGGATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG	1440			
Db	9686	CTGTCTTTTAAACTGGGATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG	9745			
QY	1441	ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTACCGGACGAGGCTGGCCT	1500			
Db	9746	ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTACCGGACGAGGCTGGCCT	9805			
QY	1501	C 1501				
Db	9806	C 9806				
RESULT 9						
AX409637	AX409637	Sequence 2284 from Patent WO0229103.	13957 bp	DNA	linear	PAT 14-JUN-2002
LOCUS	AX409637					
DEFINITION	AX409637					
ACCESSION	AX409637					
VERSION	AX409637.1	GI:21442342				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1					
AUTHORS	Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.					
TITLE	Gene expression profiles in liver cancer					
JOURNAL	Parent: WO 0229103-A 2284 11-APR-2002;					
FEATURES	GENE LOGIC INC (US)					
source	Location/Qualifiers					
	1. .13957					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					
	/note="EMBL/GenBank Accession No. M18533"					
ORIGIN						
Query Match	100.0%;	Score 1501;	DB 6;	Length 13957;		
Best Local Similarity	100.0%;	Pred. No. 0;				
Matches 1501;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
QY	1	ACAAATGGCAAGACCTTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAACCTGGA	60			
Db	8416	ACAAATGGCAAGACCTTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAACCTGGA	8475			
QY	61	TGAAAAAGCAAAAAAATCTTGAGATCCCTGGAAAGGTTCCGATGATGAGTCTCTGTTACA	120			
Db	8476	TGAAAAAGCAAAAAAATCTTGAGATCCCTGGAAAGGTTCCGATGATGAGTCTCTGTTACA	8535			
QY	121	AAGACCTTTGATATACATGAACCTTCAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACAT	180			

8536 AAGCGTTTGGATTAACATGAACCTTCAAGTGGAGTGAACCTTGGAAAAAGTCTCTCAACAT 8595
181 TAGTCCCATTTGGAAGCCAGTTCTGACCAAGTGGAGCGGTCTGCACCTTTCTCTCGAGGA 240
8596 TAGTCCCATTTGGAAGCCAGTTCTGACCAAGTGGAGCGGTCTGCACCTTTCTCTCGAGGA 8655
241 ACTTCTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGGCACTATTGGAGG 300
8656 ACTTCTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGGCACTATTGGAGG 8715
301 CGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCTTTCAAGAGGGAATTGAA 360
8716 CGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCTTTCAAGAGGGAATTGAA 8775
361 AACTAAGACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 420
8776 AACTAAGACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 8835
421 GCCTTTGGAAGGACTAGAGAACTTACAGAGAGCCAGAGAGCTGCCTCTCGAGGAGAG 480
8836 GCCTTTGGAAGGACTAGAGAACTTACAGAGAGCCAGAGAGCTGCCTCTCGAGGAGAG 8895
481 AGCCGAGATGTCACCTCGGCTTCTAGCAAGCAGGCTGAGAGGTCATATCTGAGTGGGA 540
8896 AGCCGAGATGTCACCTCGGCTTCTAGCAAGCAGGCTGAGAGGTCATATCTGAGTGGGA 8955
541 AAAATTGAACCTGCACTCCGCTGACTGCGAGAGAAAATAGATGAGACCTTTGAAGACT 600
8956 AAAATTGAACCTGCACTCCGCTGACTGCGAGAGAAAATAGATGAGACCTTTGAAGACT 9015
601 CAGGAACCTTCAAGAGCCACGAGTGAAGTGAACCTCAAGCTGCGCCCAAGCTGAGGTGAT 660
9016 CAGGAACCTTCAAGAGCCACGAGTGAAGTGAACCTCAAGCTGCGCCCAAGCTGAGGTGAT 9075
661 CAGGAGATCTGCGACGCGTGGGCGATCTCTCATTTGAGTCTCTCCAGATCACTCGA 720
9076 CAGGAGATCTGCGACGCGTGGGCGATCTCTCATTTGAGTCTCTCCAGATCACTCGA 9135
721 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCTCTGAAAGAGAAAGCTGAGCCAGTCAA 780
9136 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCTCTGAAAGAGAAAGCTGAGCCAGTCAA 9195
781 TGACCTTGCTCGCAGCTTACACTTTGGGATTCAGCTCTCAGCTATACCGTATACCTCAGAC 840
9196 TGACCTTGCTCGCAGCTTACACTTTGGGATTCAGCTCTCAGCTATACCGTATACCTCAGAC 9255
841 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAG 900
9256 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAG 9315
901 GCAGCTGCATGAAGCCCAAGGACTTTGGTCCAGCATCTCAGACATTTCTTCCAGCTC 960
9316 GCAGCTGCATGAAGCCCAAGGACTTTGGTCCAGCATCTCAGACATTTCTTCCAGCTC 9375
961 TGTCCAGGCTCCTCGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 1020
9376 TGTCCAGGCTCCTCGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 9435
1021 CGAGACTCAAACTTGTGGGACCATCCCAAAATGACAGAGCTTACCAAGTCTTTAGC 1080
9436 CGAGACTCAAACTTGTGGGACCATCCCAAAATGACAGAGCTTACCAAGTCTTTAGC 9495
1081 TCACCTGTAATATGTCAGATCTCAGCTTATAGGCTGCGCATGAACCTCCGAAAGCTGCA 1140
9496 TCACCTGTAATATGTCAGATCTCAGCTTATAGGCTGCGCATGAACCTCCGAAAGCTGCA 9555
1141 GAAGCCCTTTGCTTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCTTTGGACCA 1200
9556 GAAGCCCTTTGCTTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCTTTGGACCA 9615
1201 CAACCTCAAGCAAAATGACAGCCCATGGATATCTGCAGATTTATTAATGTTTGGACCA 1260

9616 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTTAATTTGTTTGACCAC 9675
1261 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGCTCTCTCTCGGTGGA 1320
9676 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGCTCTCTCTCGGTGGA 9735
1321 TATGCTGCTGAACGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
9736 TATGCTGCTGAACGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 9795
1381 CCTGCTTTTAAAACTGGCATCATTTCCCTGCTGTAAGACATTTGGAGCAAGTACAG 1440
9796 CCTGCTTTTAAAACTGGCATCATTTCCCTGCTGTAAGACATTTGGAGCAAGTACAG 9855
1441 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACAGGCGAGGCTGGCCT 1500
9856 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACAGGCGAGGCTGGCCT 9915
1501 C 1501
9916 C 9916

RESULT 10
AX538581
LOCUS
DEFINITION Sequence 1 from Patent WO0229056.
ACCESSION AX538581
VERSION AX538581.1 GI:25271086
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Chamberlain, J.S. and Harper, S.O.
Mini-dystrophin nucleic acid and peptide sequences
Patent: WO 0229056-A 1 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
source 1..13957
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 1501; DB 6; Length 13957;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAATGGCAAGACCTTCAAGGTGAAATTTGAAGCTCACAGATGTTTATCAACCTGGA 60
DB 8416 ACAATGGCAAGACCTTCAAGGTGAAATTTGAAGCTCACAGATGTTTATCAACCTGGA 8475
QY 61 TGAARACACCCAAAATCCTGAGATCCCTGAAAGTTCGAGTATGAGTCTCTTACA 120
DB 8476 TGAARACACCCAAAATCCTGAGATCCCTGAAAGTTCGAGTATGAGTCTCTTACA 8535
QY 121 AAGACGTTTGGTAAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACAT 180
DB 8536 AAGACGTTTGGTAAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACAT 8595
QY 181 TAGGTCCCAATTTGGAAGCCAGTTCTGACCAAGTGGAGGCTGACACCTTTCTCTCAGGA 240
DB 8596 TAGGTCCCAATTTGGAAGCCAGTTCTGACCAAGTGGAGGCTGACACCTTTCTCTCAGGA 8655
QY 241 ACTTCTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGGCACTATTGGAGG 300
DB 8656 ACTTCTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGGCACTATTGGAGG 8715
QY 301 CGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCTTTCAAGAGGGAATTGAA 360
DB 8716 CGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCTTTCAAGAGGGAATTGAA 8775

QY	1441	ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGGAGCTGGGCT	1500
Db	9856	ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGGAGCTGGGCT	9915
QY	1501	C 1501	
Db	9916	C 9916	
RESULT 11			
HUMDYS			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
COMMENT			
FEATURES			
source			
gene			
CDS			

13957 bp mRNA linear
Homo sapiens dystrophin (DMD) mRNA, complete cds.
M18533 M17154 M18026 M20250
M18533.1 GI:181856
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1699)
Koenig,M., Hoffman,E.P., Bertelson,C.J., Monaco,A.P., Feener,C. and
Kunkel,L.M.
Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and
preliminary genomic organization of the DMD gene in normal and
affected individuals
Cell 50 (3), 509-517 (1987)
87273512
3607877
2 (bases 1678 to 3830)
Hoffman,E.P., Monaco,A.P., Feener,C.C. and Kunkel,L.M.
Conservation of the Duchenne muscular dystrophy gene in mice and
humans
Science 238 (4825), 347-350 (1987)
88018015
3659917
3 (bases 1 to 13957)
Koenig,M., Monaco,A.P. and Kunkel,L.M.
The complete sequence of dystrophin predicts a rod-shaped
cytoskeletal protein
Cell 53 (2), 219-226 (1988)
88194521
3282674
On May 25, 2000 this sequence version replaced gi:340693.
Draft entry and computer-readable sequence kindly provided by
M. Koenig, 01-APR-1988 The severity of muscular dystrophy is
determined by the size of the deleted DNA segment. Deletions found
in different patients were from positions 302-2200, 473-1168,
1691-1810, and 1169-3011.
Location/Qualifiers
1..13957
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp21.3-p21.1"
/tissue_type="muscle"
/dev_stage="fetus"
/note="G00-119-850"
1..13957
/gene="DMD"
209..11266
/gene="DMD"
/codon_start=1
/product="dystrophin"
/protein_id="AA53189.1"
/db_xref="GI:181857"
/translation="MLWEEVEDCYEREDVQKFTTKWNAQFSKFGKHQIENLFSDL
QDGRLLDLLEGITGQKLPKKGSTRVHALLNNVKNALRVLNNVLDVNIIGSTDI VDG
NHKLTGLIWNIIILHWQVKNMKNMAGLQCTNSEKILLSWRQSTNYQVNVINFT
TSWSDGALNALIHSHRPDLFDWNSVCCQSATORLEHARNIARYQIGIKELDPEDV
DTTPDKKSILMYITSLFQVLPQOVSIEAIQEVEMPRPKVTKKEHFQLHQMHIQ
QITVSLAQGYERTSSPKPRFKSYATQAAVYTTSDPTRESPPFQKLEAPDKPGSSL


```

MESEVNLDRVQTALEBVLWLLSABDTLQAQGEISNDVEVVKDQPHTHEGYMDLTAH
QGRVNILOGSKLIGTKLSEDETEVQEMNLNSRWECRLVASMEQSNLHVLAM
DLQNKILDELNLTKTEERTKMEEPGLPDLKRLQVQKHVLEQDLBOEVRVN
SLTHMVVSESSDHATAALEBOLVILGRWANI CRWTEDEWLLDILLAKWRLTE
EOCLFSAWLSEDESDAVNKIHTTGFQDNEMLSLOKLAVLKADLEKKQSGMKLSLK
ODLLSLKNSVTKTEAMLDNPARCNDNLVQLKSKSTAQISQAVTTQPSLTQITVM
ETVTVTTBQILLKHAQELPPFPKKRQITVDSIEIRKLDVDTILHLSWITSEEA
VLQSPFAIFRGENSDLEKNAIEREKAERKFLQDASRAQALVQWMEVNEGA
DSIKQASELNSWIEFCOLLGERLNLWLEYQNNI IAFYNQLQQLBQMVTITAEWLKIQ
PTTSPPTAIKSKLICKDEKVNRLSGLQPIERLKTISALKKEGQPMFLDADVAF
THNFKQVFDVOAREXELOT I PDLTPPMYQRTMSAI RTWVQOSTKLISIPOLSVTDY
EIMEORLQALQSSLOQSSGLVYLSTTVKEMSKAPSEISRSQSEFFIEGRWK
KLSSQLVHCQKLEOMNKLRLQNHQITLKWMAEVDVFLAKEBPAFGDSLEILKKQL
KQCRLLVSDIQTIPSJNSVNEGQKIKNEAPEFASRLUETELKELNTQWDMHCQVY
ARKEALUGGLKTVSLQDLSSEHMTQABEYLERDFEYKTPDELQKAVEEMKRAK
EAAQKEAKVLLTESVNSVIAQAPVAQALKELETITTYNQLCRLNAGKCKTLE
EYMACWHELLSYLEKANKLNEVEFKLKTENIPGABEISVLDLSLNLNMRHSNDP
NQTIRIAQTLTDCGVNDELNEELTFNSRWELHEEA VRROKLLLEOSIQSAQETKS
LHLIOESLTFIDKLAAYTADKVDRAQMPQEAQKIOSDLTSEHISLEEKXNQSEA
AORVLSQIDVAQKLDVSMKPRFLQKPAFELRLQESKMILDEVMHLPALETKSV
QEVVQSLNHCNVLKSLSEVSEVEMVTKGTQVQKQNTENPKELDERVTAALKHY
NEIVGATBERKQLEKCLSRMRKEMNVLTEWLAATDMELTKRSAYEGMPSNLJSE
VANGKATQKEIEKQVHLKSIITVGEALKTVLGEKTELVEDKLSLNSNWTAVTSRAE
EWNLALLEYQKMETPDQVNDHITKWIIOADTLDESEKKKPOQKEDVLRKLAELND
IRPKVDSIRDOQALMANRGRHCRKLVPOI SELNHRPAALSHRIKTGKASIPUKELE
OFNSDIQKLEPLEARIQOQVNLKSEDEKMNEDNEGTVEKELLQGNLQORTDER
KREBIKIQOLLQTKNALKDLRSRRKALBISHQWYQKQADDLKCLDDIEKKL
ASLPSRBRKKKEIDRELQKKEELNVAHQEGLSBDGAAMEVPTQIQLSKRWE
IESKFAQERLAFQITREEMVTMTDMPLEISYVFTSLTEITVQSALLEVEQ
LNLAPDLCADEFDFKQESLUNKDSLOQSSGRIDIHSHKTAALQASLPVBRVKL
QEALSDLPQKWEKVMKRYQROGRFDRSEVKRPHYDIKTFNWLTEABOFLRKTQI
PENWEHAKVYKLELQDQIGOROTVRLNATGEEILCOQSKTSDASILOKLSLNL
RWQVCNKLDRKRLBEKXNLSFQRLNEFVYWLBEADNIASIPLEPKQOLKE
KLEOVKLVLEPLRQGLKUNLETGPPVLSAPISPEBQKLEKNGKQTLNQLKVS
RALPEKQIEAQIDQGLQKLEJEDLEQNLHLLWLSPIRNGLEIYNQPNQBPFD
VQETBIAQVQKQDEBLESQHLKYEKPATQVYRKLDELSSEWKAVNRLLQSLRA
KQPDMLSLDQIGSPQTVTLVTPQVVTETATSKLEMPSLMLEVPAALADFNRAWT
ELTDLWLSLQVQKQVWVGLDENIMI KOKATMODLEQRRPOLBELITAAQNLK
NXTSNOEARTIITDRERIONOWDEVOHLNQRQOLNEMLKDSTOWLEKEEAQVL
GOARAKLESWKSGPYVDAIOKJITETQLAKOLRQWQNTVDVANDLALKLLRYSAD
DTEKVMHITENINASWRSIHKVYSEREALEBETHRLLOQPLDLEKFLAWLLEAETTA
NVLQATRLNDRNLFWSLEKXNLSIRHLEASDQWKRLLSLQELLVWLKDDDEL
SQAQPIGDFPAQVQNDVHRAFKELKTEPVMSTLSTVRIETLEPQLEKLEKVO
EPRELPPBERAQVNRLLRKOAEVNTWEKLNLSADWQRIKIDETLEPQLEKLEOYD
EUDLKLQAEVITKGSQWPGVDLILDSLOHLEKVKALRGEIAPLKENVSHVNDLAROL
TTLGQLSPYNLSTLEDLNTWRKLLQVAVDRVQLHEARDRFGPASHFLSTSVQGP
WERALSPKVPYINHEQTQTCWDFRMEIYQSLADLNVRFRFAMKRLRLQKA
LCGLLUSACALDQNLKNDQNDPMDILOIINCLFTTYDRLEQHNHNLVNPCLVD
MCLNLLWNYDTRGRIRVLSFKGTIISLCAHLEDKRYLFRQVANSSTGFCQDRL
GULLHDSQIPQLQEVASFGSNIETSPRSQFQANNKPEIEAALFDWMRLPEQSM
VMLPVLHRAAAETAHQAOKNICKEPIIGFRYSRLKHENYDICQSFQSVRAVAKH
KXHYPMVEYCTPTTSGEDVRPAKLVNKERTKEVFAKHPMGLVLPQVTLVLEGNMET
PVTLLNFWPVSAPASSPOLSHDDTHSRIEYASRLABMENSNGSYLNDISIPNESID
DEHLLIQHYQSLNODSLQPRSPAQLISLESEBERGELERLADLEENNRNQABY
DRKQOQHKGSLPSPPEMPTSPQSPRDAELIABAKLLRQHKHLEARMQILEDH
NKQLESQHLRQLLEQPOAEKAVNGTIVSSPSTSLQRSDSQPMPLRVVQSTSDSM
GBEDLLSPQDSTSLGEEVMEQLNNSFPSSRGRNTPGKPMREDTMT"

```

ORIGIN

```

Query Match      100.0%; Score 1501; DB 9; Length 13957;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ACAATGGCAAGACCTCCAAAGTGAATTAAGAGCTCACACAGATGTTTATCAACACCTTGA 60
DB      8416 ACAATGGCAAGACCTCCAAAGTGAATTAAGAGCTCACACAGATGTTTATCAACACCTTGA 8475

QY      61  TGAACAACGCCAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCCTGTTACA 120
DB      8476 TGAACAACGCCAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCCTGTTACA 8535

QY      121  AAGACGTTTGGTATAACATCAACTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACAT 180

```


Db 9616 CAACCTCAAGCAAAATGACCCAGCCATGGATATCTCGAGATTATTAATTTGTTGACCAC 9675

Qy 1261 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTCGGTGA 1320

Db 9676 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTCGGTGA 9735

Qy 1321 TATGTCCTGAACCTGGCTGCTGAATGTTTATGATACGGAGCAAGAGGAGATCCGCTGT 1380

Db 9736 TATGTCCTGAACCTGGCTGCTGAATGTTTATGATACGGAGCAAGAGGAGATCCGCTGT 9795

Qy 1381 CCGTGTCTTTTAAATCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAACAAGTACAG 1440

Db 9796 CCGTGTCTTTTAAATCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAACAAGTACAG 9855

Qy 1441 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACCCAGCCAGCTGGGCT 1500

Db 9856 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACCCAGCCAGCTGGGCT 9915

Qy 1501 C 1501

Db 9916 C 9916

RESULT 12

AR220819

LOCUS AR220819 13977 bp DNA linear PAT 26-SEP-2002

DEFINITION Sequence 60 from patent US 6426186.

ACCESSION AR220819

VERSION AR220819.1 GI:23327696

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 13977)

AUTHORS Jones,K.A., Volkuth,W. and Walker,M.G.

TITLE Bone remodeling genes

JOURNAL Patent: US 6426186-A 60 30-JUL-2002;

FEATURES

Location/Qualifiers

1..13977

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 99.8%; Score 1497.4; DB 6; Length 13977;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1498; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACATGGCAAGACCTCAAGGTGAATTTGAAGTCAAGTCAAGATGTTTATCAACCTGGA 60

Db 8416 ACAATGGCAAGACCTCAAGGTGAATTTGAAGTCAAGTCAAGATGTTTATCAACCTGGA 8475

Qy 61 TGAACACAGCAAAATCTGAGATCCCTGGAGGTTCCGATGATCGAGTCTCTGTACA 120

Db 8476 TGAACACAGCAAAATCTGAGATCCCTGGAGGTTCCGATGATCGAGTCTCTGTACA 8535

Qy 121 AAGACGTTTGATTAACATGAATTTCAAGTGGAGTGAATTCGGAAGAAAGTCTCTCAACAT 180

Db 8536 AAGACGTTTGATTAACATGAATTTCAAGTGGAGTGAATTCGGAAGAAAGTCTCTCAACAT 8595

Qy 181 TAGGTCCATTTGGAGCCAGTCTGACCAAGTGGAGGCTCTGACCTTTCTCTGAGGA 240

Db 8596 TAGGTCCATTTGGAGCCAGTCTGACCAAGTGGAGGCTCTGACCTTTCTCTGAGGA 8655

Qy 241 ACTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGCCGAGCAGCACTATTGGAGG 300

Db 8656 ACTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGCCGAGCAGCACTATTGGAGG 8715

Qy 301 CGACTTTCCAGCAGTTCAAGAGCAAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 360

Db 8716 CGACTTTCCAGCAGTTCAAGAGCAAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 8775

Qy 361 AACTAAAGAACCTGTAATCATAGTACTTTGAGACTGTACGATATTTCTGACAGCA 420

Db 8776 AACTAAAGAACCTGTAATCATAGTACTCTTGAGACTGTACGAATATTTCTGACAGCA 8835

Qy 421 GCCTTTGGAGACTAGAGAACTCTACAGAGCCACAGAGGCTCCCTCTGAGGAGAG 480

Db 8836 GCCTTTGGAGACTAGAGAACTCTACAGAGCCACAGAGGCTCCCTCTGAGGAGAG 8895

Qy 481 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACCTGAGTGGGA 540

Db 8896 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACCTGAGTGGGA 8955

Qy 541 AAAATTGAACCTCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACT 600

Db 8956 AAAATTGAACCTCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACT 9015

Qy 601 CCAGGAATTTCAAGAGCCACCGATGAGCTTGGACCTCAAGCTCGCCCAAGCTGAGTGTAT 660

Db 9016 CCAGGAATTTCAAGAGCCACCGATGAGCTTGGACCTCAAGCTCGCCCAAGCTGAGTGTAT 9075

Qy 661 CAAGGAGTCTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCAAGATCACTCTGA 720

Db 9076 CAAGGAGTCTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCAAGATCACTCTGA 9135

Qy 721 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACTGAGCCACGTCAA 780

Db 9136 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACTGAGCCACGTCAA 9195

Qy 781 TGACCTTCTCGCAGCTTACCACTTTGGGCACTCAAGCTCTCAAGCTTAACTCAGCAC 840

Db 9196 TGACCTTCTCGCAGCTTACCACTTTGGGCACTCAAGCTCTCAAGCTTAACTCAGCAC 9255

Qy 841 TCTGGAAGACTGAACACACAGATGGAGGCTCTGCAAGTGGCGCTCGAGGACCGAGTCAG 900

Db 9256 TCTGGAAGACTGAACACACAGATGGAGGCTCTGCAAGTGGCGCTCGAGGACCGAGTCAG 9315

Qy 901 GCAGCTGCATGAAGCCACACAGGAGCTTTGGTCCAGCATCTCAGACATTTCTTCCACGTC 960

Db 9316 GCAGCTGCATGAAGCCACACAGGAGCTTTGGTCCAGCATCTCAGACATTTCTTCCACGTC 9375

Qy 961 TGTCCAGGGTCCCTGGGAGAGGCGCTCTCGCCAAACAAAGTGCCTTACTATACACCA 1020

Db 9376 TGTCCAGGGTCCCTGGGAGAGGCGCTCTCGCCAAACAAAGTGCCTTACTATACACCA 9435

Qy 1021 CGAGACTCAAAACAATTTGCTGGGACCATCCCAAAATGACAGAGCTCTTACGAGCTTTTAGC 1080

Db 9436 CGAGACTCAAAACAATTTGCTGGGACCATCCCAAAATGACAGAGCTCTTACGAGCTTTTAGC 9495

Qy 1081 TGACCTGAATTAATGTCAAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAGACTGCA 1140

Db 9496 TGACCTGAATTAATGTCAAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAGACTGCA 9555

Qy 1141 GAAGGCCCTTTGCTTGGATCTCTTGAAGCTGTCAGCTGTCATGTGATGCTTGGACCCAGCA 1200

Db 9556 GAAGGCCCTTTGCTTGGATCTCTTGAAGCTGTCAGCTGTCATGTGATGCTTGGACCCAGCA 9615

Qy 1201 CAACCTCAAGCAAAATGACCCAGCCCATGGATATCTCTCAGATTAATTAATTTGTTGACCA 1260

Db 9616 CAACCTCAAGCAAAATGACCCAGCCCATGGATATCTCTCAGATTAATTAATTTGTTGACCA 9675

Qy 1261 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTCGGTGGA 1320

Db 9676 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTCGGTGGA 9735

Qy 1321 TATGTCCTGAACCTGGCTGCTGAATTTATGATACGGGACGAAACAGGAGGATCCGCTGT 1380

Db 9736 TATGTCCTGAACCTGGCTGCTGAATTTATGATACGGGACGAAACAGGAGGATCCGCTGT 9795

Qy 1381 CCGTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAACAAGTACAG 1440

Db 9796 CCGTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAACAAGTACAG 9855

Qy 1441 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTACCCAGCCAGCTGGGCT 1499

Db 9856 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTACCCAGCCAGCTGGGCT 9914

RESULT 13

AF070485 13887 bp mRNA linear MAM 09-DEC-1998

LOCUS

Canis familiaris dystrophin mRNA, complete cds.

AF070485

AF070485.1 GI:3982750

Canis familiaris (dog)

Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

1 (bases 1 to 13887)

Carville, K.S., Mann, C.J., Schatzberg, S.J. and Wilton, S.D.

Direct Submission

Submitted (04-JUN-1998) ANRI, Pathology, University of Western Australia, Vardon Street, Nedlands, WA 6018, Australia

Location/Qualifiers

1. 13887

/organism="Canis familiaris"

/mol_type="mRNA"

/strain="bred Golden retriever"

/db_xref="taxon:9615"

<1..289

290..11332

/codon_start=1

/product="dystrophin"

/protein_id="AAC83646.1"

/db_xref="GI:3982751"

/translation="MLWMEVEVCYEREDVQKKTFTKWNAQPSFGKOHIEINLFSDDLQDRRLDLLEGITGQKLPEKGSRTVRHALNNKALRVLQKNVDVLNIGSTDIVDGNKHLTGLLWNILHWQKNVKNIMAGCOOTNSKILLSWRSQSTRNYPQVNVINFTTSWSDGLANALIHSHRPDLFDWNSVQCSAQRLRHAFNIAKQVLEKGLDDEPDVATYDPDKSILMTITSLFVLQPVQVSEIAIQEVMLEPRPSQVTRREHFQHHQMSYQITVLSAQGRAPSPKPRFKSYATQAAVTITSDTPRSLPSQHLTEPTSGFGRSGLTEANSDSYTALSEVLSWLSADALQAQCEISNDVEEKEQPHTEGYMMDLTHGVRNQLQASQITIGTGLSEDBETEVEQVNLNLSWECRLVASMKEQLNHLKVLMLQNLQKELNDMLTKTEBTRKMEKEPLGPIEDLRQVOQHVKLQEDLSQEVORVYNLTHMVVVVDSSGDHATAALEQLKVLGDRWANIQRWTERDVLWLDLILKWQRFESQCLPSLSEKEDAVNKHITGFKQDQSEVLSNQLKALVLTDEKKKQMDKGLSLNODILSALKNTVVAHKNMEALNDFAPQMDNLVQKLEKSAQISQAVITITQPSLTQITVMEITVITTRHILVKAQDELPPPPKQRQIIVDSEIRKRLDVIDTELHSMITRSEAVLQGEFAYIKRGNFSLDKBKVANAIREKAEFRKLQDASQALVEQVMNVEGVADISIQEQLNSRWIEFQLSERLNLWIEYQNNITFYNLQQLQEQMTTAAENLWKTQDITTSAPTSALQKICDEINRLSALQPOEIRLQISIALKQKQGMFLDADFVAFTNHFQRFADVAQAEKELQITFDLSLPPMYQETWSTLTITWIQOSETKLSIQVTVTVEYIMQRFQELQALQSSIQEQONGLYLSTTVKEMSKAPLSDIRKYSQSPREEGRWKLASQLVHDCQKLEEQMAKRIQNHIKTLKWIITVDVFLKEENFALGDSEILKRCLOQRLLIANDITQTPSLSNVGEAQKMKNEPPEFAGRLRELTRELUNTDQWYEMCQRYARKEALKGGDILTSLOKUSEHNMWTAQEEBETLRETFEYKTPDELQTAVERKQALKEEAQCKEAKVLLTSVNSVIAQAPPAQALKKELDTLTNTYQLCTLRNGKCTLEEVWAKWHELLSYLEKANKWLSEVEVKLTNTENSGAAETAEVLDLENLMQSHEDNQRILIAQTLTDGGMDELINESLETFSNRWELHEEAVRRQKLQSQSAQOIEKSLHILQPSLSDIQDLAAYTADKYDAQAPQEAQKIQSDLTSHSILEMKKHQNGLETAQRVLSQIDVAQKLODVMKPRLFQKPFANFEORLOESKMWLDVKWMLPALETKSEVQEVQSLNHCNLYKSLSEVSEVMVITKGROIYQKQTEPKELDERVATKLLHYNELGAVTERKQOEXKLVKSRKEMNALTEWLAATDMELTKRAVGEPMNSLDSEVANGKATQKEIEAQKHVLSKSTVEGALKVTJGKKEMLVEDKLSLNSLVNIAVTSRAEWNLNLTLEQGMTEFDQNVYITNNIIQADALLDESEKKKPOKEDIILKRLKAEANDIRPKVSDTQLEPAAANWANGDHCRKVVEPKISELNHRFAAISHRITKGASIFLKELEFNQSDTKLLEPALEIQOQNLKEEDFNKMSDENEGTVKELLQRGDNLQQRITDEKRKEBIKIQOLLQTKNALKDSQRKKALEISHOWYOVRQADALLCCLLDIEKKLASLPEPRDERKIKEIDRELQKKEELNAVROAEGSEDAQMAVEPTQILQSKRWREIESKPAQFRLNFPQIITHVEESVAMTDMPLIEIYVPSVTYLTETHYSQALSEVBELLNAPDLCAQDEPDLEKQESLKNIKDSLQIQISGRDIIHNKKTAAALHATPAERAQLQENALSRDLQFQWRVNNYKDRQRFDRSEKVRFRHYDMKILNQLWITEAEQFUKTKIQIPENWEHAKYKWLKELODQIGQSVVRVNLNATGEBEIIQOSSKTDASILQEKGLSLNRMQVECVQAKERKKEDEKQNLISEFDVRNBEFVLEEDANNVNTPLPNEGQQLKEKVLQKLLAAEPFLRGQIKQLNETGTGLVPSAPLSEPEQDKLENKLNQMLQWIKVSNLPEKQOIEIAHQGLQGLQELNHLHLLWSLPINQLEIYNQPNQGTGFDIKEIEVAVQAQKPDVBSGLQKQVKEFATQAKRKLIEDSSDMKVTYLLQLEQRAKQPGPAPGHTVTRAPPSQVTYITQTPAVNTKATSKIEKSEVLSIEVPLANLWATHTYDQK

5'UTR

CDS

FEATURES

source

1082 CCAGGAGCTTCAGGAGCAACAGATGAGTGGATCTCAAACTACGTCAGGAGGTCAT 9141
661 CAAGGATCTCGGAGCCCGTGGCGATCTCTCTATGACTCTCTCCAGATACCTTGA 720
9142 CAAGGATCTCGGAGCCCGTGGCGATCTCTCTATGACTCTCTCCAGATACCTTGA 9201
721 GAAAGTCAAGGACCTCGAGGAGAAATTCGGCTCTGAAAGAGACGAGGACGTCGA 780
9202 AAAAGTCAAGGCGCTTCGAGGAGAAATTCACCTCTGAAAGAGATGTCAGTACGTCGA 9261
781 TGACCTTGCTCGGAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 840
9262 TGACCTTGCTCGGAGCTTACCACTTTGGGCAATTCAGCTCTCACCATATAACCTCAGCAC 9321
841 TCTGGAAGACCTGAACACAGATGGAAGCTTCTGAGTGGCGCTCGAGGACCGAGTCAG 900
9322 TCTGGAAGACCTGAACACAGATGGAAGCTTCTGAGTGGCGCTCGAGGACCGAGTCAG 9381
901 CGAGCTGCAATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCACTTCTTTTCCAGCTC 960
9382 CGAGCTGCAATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCACTTCTTTTCCAGCTC 9441
961 TGTCCAGGTCCTCGGAGAGAGCCATCTCGGCAAAAGTGGCCCTACTATATCAACCA 1020
9442 TGTCCAGGTCCTCGGAGAGAGCCATCTCAGCAAAAGTGGCCCTACTATATCAACCA 9501
1021 CGAGCTCAAAAGCTTGGGAGCCATCCCAAAATGACAGAGCTCTACAGCTCTTTAGC 1080
9502 CGAGCCCAAAAGCTTGGGAGCCATCCCAAAATGACAGAGCTCTACAGCTCTTTAGC 9561
1081 TGACCTGAATAATGTGAGATCTTCACTTATAGGACTGCCATGAATCCGAGAGCTGA 1140
9562 TGACCTGAATAATGTGAGATCTTCACTTATAGGACTGCCATGAATCCGAGAGCTGA 9621
1141 GAAGGCCCTTGTGATCTCTGAGCTCTGAGCTGTCAGCTGATGATGCTGAGCCCTCTGCGTGA 1200
9622 GAAGGCCCTTGTGATCTCTGAGCTCTGAGCTGTCAGCTGATGATGCTGAGCCCTCTGCGTGA 9681
1201 CAACCTCAAGCAAAATGACAGAGCCATGGATATCTGAGATTAATTTGTTGACCCAC 1260
9682 CAACCTCAAGCAAAATGACAGAGCCATGGATATCTGAGATTAATTTGTTGACCCAC 9741
1261 TATTTATGAGCCGCTGGAGAGAGACAAATTTGGTCAAGCTCCCTCTCTGCGTGA 1320
9742 TATTTATGAGCCGCTGGAGAGAGACAAATTTGGTCAAGCTCCCTCTCTGCGTGA 9801
1321 TATGCTGCACTGCTGCTGATGTTTATGATACGGAGCAACAGGAGATCCGCTGT 1380
9802 TATGCTGCTCAATTTGGTCTGCTGATGTTTATGATACGGAGCAACAGGAGATCCGCTGT 9861
1381 CCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGACATTTGGAAGACAAAGTACAG 1440
9862 CCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGACATTTGGAAGACAAAGTACAG 9921
1441 ATACCTTTTCAAGCAAGTGGAGTCAACAGATTTTGGACAGGAGGAGGCTGGGCT 1500
9922 ATACCTTTTCAAGCAAGTGGAGTCAACAGATTTTGGACAGGAGGAGGCTGGGCT 9981
1501 C 1501
9982 C 9982

RESULT 14
AX306153
LOCUS 13815 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 904 from Patent WO0188188.
ACCESSION AX306153
VERSION AX306153.1 GI:17645441
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS Ishikawa, K., Asai, S., Takahashi, Y., Nagata, T. and Ishii, Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 904 22-NOV-2001;
SCHOOL Juridical Person Nihon University (JP)
FEATURES
Location/Qualifiers
1. 13815
source
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Query Match 86.5%; Score 1298.4; DB 6; Length 13815;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
Qy 1 ACAATGCAAGACCTCAAGTGAATTTGAAGTCAACAGATGTTTATCACAACTTGA 60
Db 8407 ACCATGCAAGATCTCCAGGAGAAATTTGAATTCACAGATATCTATCACAACTTGA 8466
Qy 61 TGAACACAGCAAAATCTCGATCCCTGGAAGTTCCGATGATGATGATGATGATGATGAT 120
Db 8467 TGAACATGCAAAATCTCGATCCCTGGAAGTTCCGATGATGATGATGATGATGATGAT 8526
Qy 121 AAGAGCTTGGATTAACATGAATTTCAAGTGAAGTGAATTTGGAAGAAAGTCTCTCAACAT 180
Db 8527 AAGAGCTTGGATTAACATGAATTTCAAGTGAAGTGAATTTGGAAGAAAGTCTCTCAACAT 8586
Qy 181 TAGTCCCATTTGGAAGCAAGTCTGACAGTGAAGCGTCTGACACCTTTCTCTGACGGA 240
Db 8587 TAGTCCCATTTGGAAGCAAGTCTGACAGTGAAGCGTCTGACATCTTTCTCTTCAGGA 8646
Qy 241 ACTTCTGCTGCTCAGCTGAAAGATGATGAATTAAGCCGCGAGCAGCCTATTTGAGG 300
Db 8647 ACTTCTGCTGCTCAGCTGAAAGATGATGAATTAAGCCGCGAGCAGCCTATTTGAGG 8706
Qy 301 CGATTTCCAGAGCTCAGAGCAGAGATGATGATGAGGCTTCAAGAGGAAATTTGAA 360
Db 8707 TGATTTCCAGAGCTCAGAGCAGAGATGATGATGAGGCTTCAAGAGGAAATTTGAA 8766
Qy 361 AACTAAAGAACTGTATATCATGATCTCTTGAGACTGTAGAAATTTCTGACAGAGCA 420
Db 8767 AACTAAAGAACTGTATATCATGATCTCTTGAGACTGTAGAAATTTCTGACAGAGCA 8826
Qy 421 GCCTTTGGAAGCACTAGAGAACTCTACCAGGAGCCAGAGAGTGCCTCTGAGGAGAG 480
Db 8827 GCCTTTGGAAGCACTAGAGAACTCTACCAGGAGCCAGAGAGTGCCTCTGAGGAGAG 8886
Qy 481 AGCCAGAGATGTCAGTCCGCTTCTACGAAAGCAGGCTGAGGAGTCAATATCTGAGTGGGA 540
Db 8887 AGCTCAGAAATGTCAGTCCGCTTCTACGAAAGCAGGCTGAGGAGTCAATATCTGAGTGGGA 8946
Qy 541 AAAATTGAACCTGCACTCCGCTGATGCGAGAGAAATATAGATGAGACCTTTGAAAGACT 600
Db 8947 CAAATTGAACCTGCGCTCAGCTGATGCGAGAGAAATATAGATGAGACCTTTGAAAGACT 9006
Qy 601 CCAGAAATTTCAAGAGGCGAGGATGAGCTCAAGCTCAAGCTGCGCCAGGCTGAGGTCAT 660
Db 9007 CCAGAAATTTCAAGAGGCGAGGATGAGCTCAAGCTCAAGCTGCGCCAGGCTGAGGTCAT 9066
Qy 661 CAAGGAGTCTTGGCAGCCCGTGGGCGATCTCTCTCATTTGACTCTCTCCAAAGATCACTTGA 720
Db 9067 CAAGGAGTCTTGGCAGCCCGTGGGCGATCTCTCTCATTTGACTCTCTCCAAAGATCACTTGA 9126
Qy 721 GAAAGTCAAGGACCTTCAGAGAGAAATTTGCGCTTGAAGAGACGAGTGAAGCCAGTCAA 780
Db 9127 AAAAGTCAAGGACCTTCAGAGAGAAATTTGCGCTTGAAGAGATGTCATCTGTCGTC 9186
Qy 781 TGACCTTGCTGCGCAGCTTACCACTTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 840
Db 9187 TGACCTTGCTGCGCAGCTTACCACTTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 9246

```
QY 841 TCTGAAAGACCTGAACACCCAGATGGAAGCTTCTGAGGTGGCGGTGAGGACCGAGTCAG 900
Db 9247 TTTGGAAGATCTGAATACCAATGAGAGCTTCTACAGTGGCTGTGGAGGACCGTGTGAG 9306
QY 901 GCAGCTGATGAACCCACAGGAGACTTTGGTTCAGAGATCTCAGACATTTCTTTCCACGTC 960
Db 9307 ACAGCTGCATGAAGCCACAGGAGACTTTGGTCTGATCCAGCACTTCTCTTCCACTTC 9366
QY 961 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCTTACTATATCAACCA 1020
Db 9367 AGTTCAGGCTCCCTGGGAGAGAGCCATCTCACCACAAAGTGGCTTACTATATCAACCA 9426
QY 1021 CGAGACTCAAAACAACTTGTGGAGCAATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 1080
Db 9427 CGAGACCCAAACCACTTTGTGGAGCAATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 9486
QY 1081 TGACCTGAATAAGTCAAGATTTCTCAGCTTATAGCACTGCATGAAGTCCGAGAGCTGCA 1140
Db 9487 TGACCTGAATAAGTCAAGTCTCCGCTATAGCACTGCATGAAGTCCGAGAGCTTCCA 9546
QY 1141 GAAGGCCCTTTGCTTGGATCTTTGAGCCTGTGAGCTGCTGATGATGCTTGGACAGCA 1200
Db 9547 GAAGGCCCTTTGCTTGGATCTTTGAGCCTGTGAGCTGCTGATGATGCTTGGACAGCA 9606
QY 1201 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTTAAATGTTTGAACAC 1260
Db 9607 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTTAAATGTTTGAACAC 9666
QY 1261 TATTTATGACCGCTGGAGCAAGACACAAATTTGGTCAAGCTCCCTCTCTGGTGGGA 1320
Db 9667 AATTTATGATGCTGGAGCAAGACACAAATTTGGTCAAGCTCCCTCTCTGGTGGGA 9726
QY 1321 TATGTTCTGAATGCTGCTGCTGAATGTTTATGATACGGAACCAAGGAGGATCCGTG 1380
Db 9727 TATGTTCTGAATGCTGCTGCTGAATGTTTATGATACGGAACCAAGGAGGATCCGTG 9786
QY 1381 CTTGCTCTTTAAACTGGGATCATTTCCCTGTTGAAGACACATTTGGAGACCAAGTACAG 1440
Db 9787 CTTGCTCTTTAAACTGGGATCATTTCTCTGTTGAAGACACATTTGGAGACCAAGTACAG 9846
QY 1441 ATACCTTTTCAAGCAAGTGGAGTTCACAGGATTTGTGACCGGAGGCTGGGCT 1500
Db 9847 ATACCTTTTCAAGCAAGTGGAGTTCACAGGATTTGTGACCGGAGGCTGGGCT 9906

RESULT 15
AX538582
LOCUS AX538582
DEFINITION Sequence 2 from Patent WO0229056.
ACCESSION AX538582
VERSION AX538582.1 GI:25271088
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
TITLE Chamberlain, J. S. and Harper, S. Q.
JOURNAL Mini-dystrophin nucleic acid and peptide sequences
Patent: WO 0229056-A 2 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source
1. .13815
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Query Match 86.5%; Score 1298.4; DB 6; Length 13815;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
QY 1 ACAATGGCAAGACCTCCAAAGGTGAATGAAGCTCACACAGATGTTATCAACACCTGGA 60
```

Db	9487	TGACCTGAATAATGTCAGGTTCTCGCGTATAGGACTGCCATGAAGCTCAGAAGGCTCCA	9546
Qy	1141	GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCTTGGACCAAGCA	1200
Db	9547	GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCTTGGACCAAGCA	9606
Qy	1201	CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAAC	1260
Db	9607	CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACTAC	9666
Qy	1261	TATTTATGACCGCCTGGAGCAAGACCAACAATTTGGTCAACGTCCTCTCTCTGCGTGA	1320
Db	9667	AATTTATGATCGTCTGGAGCAAGACCAACAATTTGGTCAATGTCCTCTCTCTGCGTGA	9726
Qy	1321	TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGCAACAGGGAGGATCCGTTGT	1380
Db	9727	TATGTGTCTCACTGGCTTCTCAATGTTTATGATACGGGACGCAACAGGGAGGATCCGTTGT	9786
Qy	1381	CCTGTCTTTTAAACTGGCAATCATTCCCTGTGTAAAGCAATTTGGAAAGCAAGTACAG	1440
Db	9787	CCTGTCTTTTAAACTGGCAATCATTCTCTGTGTAAAGCAATTTGGAAAGCAAGTACAG	9846
Qy	1441	ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGAGGCTGGGCT	1500
Db	9847	ATACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTTGTGACCAAGCGTGGGCTCT	9906

Search completed: September 19, 2004, 08:30:01
Job time : 3980.06 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2004, 07:39:06 ; Search time 408.027 Seconds
(without alignments)
15627.746 Million cell updates/sec

Title: US-09-845-416-10_COPY_1500_3000
Perfect score: 1501
Sequence: 1 acaatgggaagacctccaag.....accagcgagctggggcctc 1501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002s: *
7: Geneseqn2003as: *
8: Geneseqn2003bs: *
9: Geneseqn2003cs: *
10: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1501	100.0	2169	6 AAD37232	Aad37232 Human dys
2	1501	100.0	3531	6 AAD37238	Aad37238 Human dys
3	1501	100.0	3858	6 AAD37237	Aad37237 Human dys
4	1501	100.0	3999	6 AAD37234	Aad37234 Human dys
5	1501	100.0	4182	6 AAD37230	Aad37230 Human dys
6	1501	100.0	4498	6 AAD37258	Aad37258 Adeno-ass
7	1501	100.0	4825	6 AAD37257	Aad37257 Adeno-ass
8	1501	100.0	4848	6 AAD37263	Aad37263 Adeno-ass
9	1501	100.0	4966	6 AAD37256	Aad37256 Adeno-ass
10	1501	100.0	4990	6 AAD37262	Aad37262 Adeno-ass
11	1501	100.0	5060	6 AAD37264	Aad37264 Adeno-ass
12	1501	100.0	5149	6 AAD37255	Aad37255 Adeno-ass
13	1501	100.0	5339	6 ABK81998	Abk81998 DNA encod
14	1501	100.0	5462	6 ABK81999	Abk81999 Human dys
15	1501	100.0	5952	5 AAD06794	Aad06794 Human dys
16	1501	100.0	8689	6 ABK82000	Abk82000 Human dys
17	1501	100.0	11058	6 AAD37229	Aad37229 Human dys
18	1501	100.0	11241	6 ABK82005	Abk82005 cDNA enco
19	1501	100.0	11443	6 ABK82002	Abk82002 DNA encod
20	1501	100.0	12923	1 AAN90338	Aan90338 Sequence
21	1501	100.0	13957	6 ABK81959	Abk81959 cDNA enco
22	1501	100.0	13957	6 ABK10904	Abt10904 Human bre
23	1501	100.0	13957	6 ABN95786	Abn95786 Gene #228

24	1501	100.0	13957	6 ABS69900	Abbs69900 Human dys
25	1497.4	99.8	13977	6 ABS70403	Abbs70403 Human bon
26	1307.8	87.1	3510	6 AAD37240	Aad37240 Human dys
27	1307.8	87.1	4476	6 AAD37259	Aad37259 Adeno-ass
28	1303	86.8	1821	6 AAD37241	Aad37241 Human dys
29	1298.4	86.5	13815	6 ABK81960	Abk81960 cDNA enco
30	1298.4	86.5	13815	6 AB199799	Abi199799 Mouse isc
31	1298.4	86.5	19307	2 AAT27558	Aat27558 Shuttle v
32	1298.4	86.5	13815	2 AAV18885	Aav18885 Mus muscu
33	985.2	65.6	4402	3 AAZ48567	Aaz48567 A rod sho
34	917.6	61.1	3446	6 AAD37242	Aad37242 Human dys
35	917.6	61.1	4414	6 AAD37260	Aad37260 Adeno-ass
36	916.6	61.1	5417	6 ABK81997	Abk81997 DNA encod
37	916	61.0	1434	6 AAD37243	Aad37243 Human dys
38	786	52.4	4402	3 AAZ48568	Aaz48568 A rod sho
39	650.4	43.3	4075	3 AAZ48569	Aaz48569 A rod sho
40	646.2	43.1	3747	3 AAZ48566	Aaz48566 A rod sho
41	619.2	41.3	10705	7 ABT41896	Abt41896 Toxicity
42	619.2	41.3	11096	6 ABK81962	Abk81962 cDNA enco
43	604.8	40.3	6045	2 AAT74665	Aat74665 Utrophin
44	604.8	40.3	6059	4 AAF84673	Aaf84673 Nucleotid
45	604.8	40.3	9195	7 ACD19399	Acd19399 cDNA enco

ALIGNMENTS

RESULT 1
AAD37232
ID AAD37232 standard; DNA; 2169 BP.

XX AC AAD37232;
XX DT 21-AUG-2002 (first entry)

XX DE Human dystrophin rod, hinge and CR domain regions encoding DNA #1.

XX KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

XX KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX KW Becker muscular dystrophy; ds.

XX OS Homo sapiens.

XX PN WO200183695-A2.

XX XX 08-NOV-2001.

XX PF 27-APR-2001; 2001WO-US013677.

XX PR 28-APR-2000; 2000US-0200777P.

XX XX (XIAO/) XIAO X.

XX PI Xiao X;

XX XX WPI; 2002-049342/06.

XX XX New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene.

XX XX Example 1; Page 45-46; 71pp; English.

XX XX The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin DNA fragment encoding rods R22, R23 and R24, hinge H4 and CR
CC domain regions
XX
SQ Sequence 2169 BP; 623 A; 529 C; 524 G; 493 T; 0 U; 0 Other;

Query Match 100.0%; Score 1501; DB 6; Length 2169;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAATGCAAGACCTCCCAAGGTGAAATTTGAAGCTTCACACAGATGTTTATCAACACCTCGA 60
DB 151 ACAATGCAAGACCTCCCAAGGTGAAATTTGAAGCTTCACACAGATGTTTATCAACACCTCGA 210

QY 61 TGAACACAGCAAAATCTGAGATCCCTTGAAGGTTCGATGATGCGAGTCTCTGTACA 120
DB 211 TGAACACAGCAAAATCTGAGATCCCTTGAAGGTTCGATGATGCGAGTCTCTGTACA 270

QY 121 AAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTCGAAAAAGTCTCTCAACAT 180
DB 271 AAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTCGAAAAAGTCTCTCAACAT 330

QY 181 TAGTCTCCATTTGGAAGCCAGTTCTGACACAGTGGAAAGCTCTGCACCTTTCTCTGCAGGA 240
DB 331 TAGTCTCCATTTGGAAGCCAGTTCTGACACAGTGGAAAGCTCTGCACCTTTCTCTGCAGGA 390

QY 241 ACTTCTGGTGTGCTACAGCTGAAGATGATGAATTAAGCCGACGACCTATTGGAGG 300
DB 391 ACTTCTGGTGTGCTACAGCTGAAGATGATGAATTAAGCCGACGACCTATTGGAGG 450

QY 301 CGACTTTCCAGCAGTTTCAGAGCAGAACGATGATAGGCTTCAAGAGGGAATTTGAA 360
DB 451 CGACTTTCCAGCAGTTTCAGAGCAGAACGATGATAGGCTTCAAGAGGGAATTTGAA 510

QY 361 AACTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATTTCTTCAACAGACA 420
DB 511 AACTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATTTCTTCAACAGACA 570

QY 421 GCCTTTGGAGGAGTACAGAACTCTACAGAGCCGACAGAGCTGCCTCTGAGGAGAG 480
DB 571 GCCTTTGGAGGAGTACAGAACTCTACAGAGCCGACAGAGCTGCCTCTGAGGAGAG 630

QY 481 AGCCCAAGATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATFACTGAGTGGGA 540
DB 631 AGCCCAAGATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATFACTGAGTGGGA 690

QY 541 ARAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 600
DB 691 ARAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 750

QY 601 CCAGGAACCTTCAAGAGGCCACGATGAGTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 660
DB 751 CCAGGAACCTTCAAGAGGCCACGATGAGTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 810

QY 661 CAAGGATCTGCGACCGCTGGGCAATCTCTTATGACTCTCTTCAAGATCACTCGA 720
DB 811 CAAGGATCTGCGACCGCTGGGCAATCTCTTATGACTCTCTTCAAGATCACTCGA 870

QY 721 GAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGCCAGTCAA 780
DB 871 GAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGCCAGTCAA 930

QY 781 TGACCTTGCTCGCAGCTTACCACTTTGGGCAATTCAGTCTCAACGTTAACTCAGCAC 840
DB 931 TGACCTTGCTCGCAGCTTACCACTTTGGGCAATTCAGTCTCAACGTTAACTCAGCAC 990

QY 841 TCTGGAAGACCTGAACACACAGATGAAGCTTCTGAGGTGGCGTGGAGCCGAGTCAG 900
DB 991 TCTGGAAGACCTGAACACACAGATGAAGCTTCTGAGGTGGCGTGGAGCCGAGTCAG 1050

QY 901 GGAGCTGATGAAGCCCAAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTC 960
DB 1051 GGAGCTGATGAAGCCCAAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTC 1110

QY 961 TGTCCAGGTCCTCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 1020
DB 1111 TGTCCAGGTCCTCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 1170

QY 1021 CGAGACTCAAAACAACCTTGTGGACCAATCCCAAAATGACAGAGCTCTTACAGTCTTTAGC 1080
DB 1171 CGAGACTCAAAACAACCTTGTGGACCAATCCCAAAATGACAGAGCTCTTACAGTCTTTAGC 1230

QY 1081 TGACCTGAATTAATCTCAGATTTCTCAGCTTATAGACTGCCATGAATCCGAGACTGCA 1140
DB 1231 TGACCTGAATTAATCTCAGATTTCTCAGCTTATAGACTGCCATGAATCCGAGACTGCA 1290

QY 1141 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGTCATGTGATGCTTGGACCAAGCA 1200
DB 1291 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGTCATGTGATGCTTGGACCAAGCA 1350

QY 1201 CAACCTCAAGCAAAATGACAGCCCAATGAGTATCTCTCAGATTAATTAATTTGTTGACCAC 1260
DB 1351 CAACCTCAAGCAAAATGACAGCCCAATGAGTATCTCTCAGATTAATTAATTTGTTGACCAC 1410

QY 1261 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGCTCAAGCTCCCTCTCTGCGTGA 1320
DB 1411 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGCTCAAGCTCCCTCTCTGCGTGA 1470

QY 1321 TATGTGCTGAACCTGGCTGCTGAATTTATGATAGCGGACGAACAGGAGGATCCGTGT 1380
DB 1471 TATGTGCTGAACCTGGCTGCTGAATTTATGATAGCGGACGAACAGGAGGATCCGTGT 1530

QY 1381 CCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAGACACATTTGGAAGACAAAGTACAG 1440
DB 1531 CCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAGACACATTTGGAAGACAAAGTACAG 1590

QY 1441 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAAGCGAGGCTGGGCT 1500
DB 1591 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAAGCGAGGCTGGGCT 1650

QY 1501 C 1501
DB 1651 C 1651

RESULT 2
AAD37238
ID AAD37238 standard; DNA; 3531 BP.
XX
AC AAD37238;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3531.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN W0200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-020077P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PI gene.

XX Example 1; Page 50-51; 71pp; English.

PS The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus,
CC hinge H1 and rod R1), 8059-10227 (rods R22, R23, R24, hinge H4 and CR
CC domain) and 11047-11058 (dystrophin last 3 amino acids)

XX Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 U; 0 Other;

Query Match 100.0%; Score 1501; DB 6; Length 3531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ACAATGGCAAGACCTCCAAAGGTGAATTTGAAGCTCACAGATGTTTATCAACCTGGA	60
DB	1500	ACAATGGCAAGACCTCCAAAGGTGAATTTGAAGCTCACAGATGTTTATCAACCTGGA	1559
QY	61	TGAAACACGCCAAAATCTGAGATCCCTGGAAGTTCGGATGATGAGTCTCTTACA	120
DB	1560	TGAAACACGCCAAAATCTGAGATCCCTGGAAGTTCGGATGATGAGTCTCTTACA	1619
QY	121	AAGACCTTTGGATAACATGAATTTCAAGTGGAGTGAATTCGGAAGTCTCTCAACAT	180
DB	1620	AAGACCTTTGGATAACATGAATTTCAAGTGGAGTGAATTCGGAAGTCTCTCAACAT	1679
QY	181	TAGGTCCCATTTGGAGGACGATTTGACAGTGGAGGCTTGACCTTTCTCTGAGAG	240
DB	1680	TAGGTCCCATTTGGAGGACGATTTGACAGTGGAGGCTTGACCTTTCTCTGAGAG	1739
QY	241	ACTTCTGGTGGCTACAGTGAAGATGATGAATTAAGCCGAGGACCTATTGGAGG	300
DB	1740	ACTTCTGGTGGCTACAGTGAAGATGATGAATTAAGCCGAGGACCTATTGGAGG	1799
QY	301	CGACTTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGA	360
DB	1800	CGACTTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGA	1859
QY	361	AACCTAAGAACCTGTAACTAGTACTCTTGAGACTGTACGAAATTTCTGACAGAGCA	420
DB	1860	AACCTAAGAACCTGTAACTAGTACTCTTGAGACTGTACGAAATTTCTGACAGAGCA	1919
QY	421	GCCTTTGGAAGGACTAGAGAACTCTACAGAGGACCCAGAGAGCTCCCTCCTGAGGAG	480
DB	1920	GCCTTTGGAAGGACTAGAGAACTCTACAGAGGACCCAGAGAGCTCCCTCCTGAGGAG	1979
QY	481	AGCCCAAGATGTCACTGGCTTTCTAGAAAGCGGTGTAGGAGGTCAATCTGAGTGGGA	540
DB	1980	AGCCCAAGATGTCACTGGCTTTCTAGAAAGCGGTGTAGGAGGTCAATCTGAGTGGGA	2039
QY	541	AAATTTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGCT	600
DB	2040	AAATTTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGCT	2099
QY	601	CCAGGAACCTTCAAGAGGCCCGGATGAGTGGACCTCAAGCTCGGCCAAGCTGAGTGAT	660
DB	2100	CCAGGAACCTTCAAGAGGCCCGGATGAGTGGACCTCAAGCTCGGCCAAGCTGAGTGAT	2159
QY	661	CAAGGGATCTGGCAGCCCTGGGCGATCTCTCATTTGACTCTCCAGATCACCTCGA	720
DB	2160	CAAGGGATCTGGCAGCCCTGGGCGATCTCTCATTTGACTCTCTCCAGATCACCTCGA	2219

QY	721	GAAGTCAAGGCACTTCGAGGAGAAATTCGCCCTCTCTGAAAGAGAACTGAGCCACTCAA	780
DB	2220	GAAGTCAAGGCACTTCGAGGAGAAATTCGCCCTCTCTGAAAGAGAACTGAGCCACTCAA	2279
QY	781	TGACCTTGCTGGCAGCTTACCACTTTGGGCACTTACAGCTCTACCGTATAACCTCAGCAC	840
DB	2280	TGACCTTGCTGGCAGCTTACCACTTTGGGCACTTACAGCTCTACCGTATAACCTCAGCAC	2339
QY	841	TTTGGAGACCTGAACACACAGATGGAAGCTTCTGCAAGTGGCCGCTCGAGGACCGAGTCAG	900
DB	2340	TTTGGAGACCTGAACACACAGATGGAAGCTTCTGCAAGTGGCCGCTCGAGGACCGAGTCAG	2399
QY	901	GCAGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGACCTTCTTCCAGCTC	960
DB	2400	GCAGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGACCTTCTTCCAGCTC	2459
QY	961	TGTCAGGGTCCCTGGGAGAGGACCTCTCGCCAAACAAAGTGCCTTACTATATCAACCA	1020
DB	2460	TGTCAGGGTCCCTGGGAGAGGACCTCTCGCCAAACAAAGTGCCTTACTATATCAACCA	2519
QY	1021	CGAGACTCAAAACAACTTTGCTGGGACCAATCCCAAAATGACAGAGCTCTACAGTCTTTAGC	1080
DB	2520	CGAGACTCAAAACAACTTTGCTGGGACCAATCCCAAAATGACAGAGCTCTACAGTCTTTAGC	2579
QY	1081	TGACCTGAATAATGTCTAGATTTCTCAGCTTATAGACTGCCATGAAACTCCGAGACTGCA	1140
DB	2580	TGACCTGAATAATGTCTAGATTTCTCAGCTTATAGACTGCCATGAAACTCCGAGACTGCA	2639
QY	1141	GAAGGCCCTTTGCTTTGGATCTCTTGAGCCTGTGATGATGATGATGATGATGATGATGATG	1200
DB	2640	GAAGGCCCTTTGCTTTGGATCTCTTGAGCCTGTGATGATGATGATGATGATGATGATGATG	2699
QY	1201	CAACCTCAAGCAAAATGACAGCCCATGGATATCTCAGATTTAATTAATTTGTTGACCA	1260
DB	2700	CAACCTCAAGCAAAATGACAGCCCATGGATATCTCAGATTTAATTAATTTGTTGACCA	2759
QY	1261	TATTTATGACCGCTGAGCAAGAGCAACAAATTTGCTCAACCTCTCTCTGCGTGA	1320
DB	2760	TATTTATGACCGCTGAGCAAGAGCAACAAATTTGCTCAACCTCTCTCTGCGTGA	2819
QY	1321	TATGCTGTGAATCTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG	1380
DB	2820	TATGCTGTGAATCTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG	2879
QY	1381	CCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAGACATTTTGAAGACAACTGATACAG	1440
DB	2880	CCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAGACATTTTGAAGACAACTGATACAG	2939
QY	1441	ATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACAGCGCAGGCTGGGCT	1500
DB	2940	ATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACAGCGCAGGCTGGGCT	2999
QY	1501	C 1501	
DB	3000	C 3000	

RESULT 3
AAD37237

ID AAD37237 standard; DNA; 3858 BP.

XX AAD37237;

XX 21-AUG-2002 (first entry)

XX Human dystrophin minigene delta3549.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX Becker muscular dystrophy; ds.

XX Homo sapiens.

XX WO200183695-A2.
 XX 08-NOV-2001.
 XX 27-APR-2001; 2001WO-US013677.
 XX 28-APR-2000; 2000US-0200777P.
 XX (XIAO/) XIAO X.
 XX Xiao X;
 XX WPI; 2002-049342/06.
 XX
 PT New dystrophin minigene for treating Duchenne or Becker muscular
 PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
 PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
 PT gene.
 XX
 PS Example 1; Page 48-49; 71pp; English.
 XX
 CC The present invention relates to an isolated nucleotide sequence encoding
 CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
 CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
 CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
 CC invention also relates to a recombinant adeno-associated virus (AAV)
 CC comprising dystrophin minigene operably linked to an expression control
 CC element. The dystrophin minigene in operable linkage with an expression
 CC control element, in a recombinant adeno-associated virus or retrovirus is
 CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
 CC dystrophy (BMD) in a mammalian subject. The present sequence is human
 CC dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
 CC hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
 CC and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
 XX
 SQ Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1501; DB 6; Length 3858;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ACAAATGGGAGACCTCAAGGTGAATTCGAAGCTCACAGATGTTATFCAAACTGGA	60
DB	1827	ACAATGGGAGACCTCAAGGTGAATTCGAAGCTCACAGATGTTATFCAAACTGGA	1886
QY	61	TGAAAACAGCCAAAATCCTGAGATCCCTGGAAGTTCGATGATGAGTCCCTGTTACA	120
DB	1887	TGAAAACAGCCAAAATCCTGAGATCCCTGGAAGTTCGATGATGAGTCCCTGTTACA	1946
QY	121	AGACGTTTGGATAACATGAATTCGAAGTGGAGTGAACCTCGAAAAAGTCTCTCAACAT	180
DB	1947	AGACGTTTGGATAACATGAATTCGAAGTGGAGTGAACCTCGAAAAAGTCTCTCAACAT	2006
QY	181	TAGTCCCAATTTGGAGCCAGTTCGACAGTGGAGCGTCTGCACTTCTCTGACGGA	240
DB	2007	TAGTCCCAATTTGGAGCCAGTTCGACAGTGGAGCGTCTGCACTTCTCTGACGGA	2066
QY	241	ACTTCGTGTGGTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG	300
DB	2067	ACTTCGTGTGGTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG	2126
QY	301	CGACTTCCAGAGTTCAGAACAGACATGATGATAGAGGCTTCAGAGGGAATTGAA	360
DB	2127	CGACTTCCAGAGTTCAGAACAGACATGATGATAGAGGCTTCAGAGGGAATTGAA	2186
QY	361	AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA	420
DB	2187	AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA	2246
QY	421	GCCTTTGGAGACTTAGAGAACTCTACAGAGGCCAGAGCTGCCTCTCTGAGGAGAG	480
DB	2247	GCCTTTGGAGACTTAGAGAACTCTACAGAGGCCAGAGCTGCCTCTCTGAGGAGAG	2306

QY	481	AGCCAGAAATGTCACTCGGGCTTCTACGAAAGCAGGCTGAGGAGTCAATATCTAGTGGGA	540
DB	2307	AGCCAGAAATGTCACTCGGGCTTCTACGAAAGCAGGCTGAGGAGTCAATATCTAGTGGGA	2366
QY	541	AAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGACT	600
DB	2367	AAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGACT	2426
QY	601	CCAGAACTTCAAGAGCCACCGATGAGCTGAGCTCAAGCTGGCCCAAGCTCAGGTGAT	660
DB	2427	CCAGAACTTCAAGAGCCACCGATGAGCTGAGCTCAAGCTGGCCCAAGCTCAGGTGAT	2486
QY	661	CAAGGATCTCTGGCAGCCCGTGGCGATCTCTCANTGACTCTCTCAAGATCACTTCGA	720
DB	2487	CAAGGATCTCTGGCAGCCCGTGGCGATCTCTCANTGACTCTCTCAAGATCACTTCGA	2546
QY	721	GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCTCTGAAAGAGAACTGAGGACCACTCAA	780
DB	2547	GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCTCTGAAAGAGAACTGAGGACCACTCAA	2606
QY	781	TGACCTTGTCTGGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCA	840
DB	2607	TGACCTTGTCTGGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCA	2666
QY	841	TCTGGAGACCTGNAACACAGATGGAAGTCTTCGAGGTGGCGCTGAGGACCGAGTCAG	900
DB	2667	TCTGGAGACCTGNAACACAGATGGAAGTCTTCGAGGTGGCGCTGAGGACCGAGTCAG	2726
QY	901	GCAGCTCATGAAGCCCAACAGGACCTTTGTCTCAGCATCTCAGCACTTTCTTTCCAGCTC	960
DB	2727	GCAGCTCATGAAGCCCAACAGGACCTTTGTCTCAGCATCTCAGCACTTTCTTTCCAGCTC	2786
QY	961	TGTCAGGGTCCCTGGGAGAGACCATCTCGCCAAACAAAGTGCCTCTATATCAACCA	1020
DB	2787	TGTCAGGGTCCCTGGGAGAGACCATCTCGCCAAACAAAGTGCCTCTATATCAACCA	2846
QY	1021	CGAGACTCAAAACAACTTGCTGGGACCATCCAAATATGACAGAGCTCTACAGTCTTTAGC	1080
DB	2847	CGAGACTCAAAACAACTTGCTGGGACCATCCAAATATGACAGAGCTCTACAGTCTTTAGC	2906
QY	1081	TGACCTGAATATGTGAGATCTCAGATTCATAGGACTGCCATGAAACTCCGAAGACTGCA	1140
DB	2907	TGACCTGAATATGTGAGATTCATAGGACTGCCATGAAACTCCGAAGACTGCA	2966
QY	1141	GAAGGCCCTTTGCTTGATCTCTTGAGCCCTGTCAGCTGATGATGATGATGATGATGATG	1200
DB	2967	GAAGGCCCTTTGCTTGATCTCTTGAGCCCTGTCAGCTGATGATGATGATGATGATGATG	3026
QY	1201	CAACCTCAAGCAAAATGACAGCCCATGGATATTCCTGCAAGATTAATAATTTGTTTGACCA	1260
DB	3027	CAACCTCAAGCAAAATGACAGCCCATGGATATTCCTGCAAGATTAATAATTTGTTTGACCA	3086
QY	1261	TATTTATGACCGCTGGAGCAGAGACCAACAATTTGGTCAAGCTCCCTCTCTGGGTGA	1320
DB	3087	TATTTATGACCGCTGGAGCAGAGACCAACAATTTGGTCAAGCTCCCTCTCTGGGTGA	3146
QY	1321	TATGTCTGTGAATGGCTGCTGAATTTTATGATCGGACCAAGAGGAGGATCCGTGT	1380
DB	3147	TATGTCTGTGAATGGCTGCTGAATTTTATGATCGGACCAAGAGGAGGATCCGTGT	3206
QY	1381	CTGTCTTTTAAACTGGGATCATTTCTCTGTGTAAAGCACAATTTGGAGAGCAAGTACAG	1440
DB	3207	CTGTCTTTTAAACTGGGATCATTTCTCTGTGTAAAGCACAATTTGGAGAGCAAGTACAG	3266
QY	1441	ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCT	1500
DB	3267	ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCT	3326
QY	1501	C 1501	
DB	3327	C 3327	

RESULT 4	QY	241	ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACTATTGGAGG	300
AD37234	DB	2208	ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACTATTGGAGG	2267
AC	QY	301	CGACTTTCCAGCAGTTTCCAGAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGGAATTGAA	360
XX	DB	2268	CGACTTTCCAGCAGTTTCCAGAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGGAATTGAA	2327
DT	QY	361	AACATAAGAACCTCTAATCATGATGATCTCTTGAGACTGTACGAATATTTCTGCACAGAGCA	420
XX	DB	2328	AACATAAGAACCTCTAATCATGATGATCTCTTGAGACTGTACGAATATTTCTGCACAGAGCA	2387
DE	QY	421	GCCTTTTGAAGGACTAGAGAAACTCTACAGAGGCCCCAGAGAGCTGCTCTCTGAGGAGAG	480
XX	DB	2388	GCCTTTTGAAGGACTAGAGAAACTCTACAGAGGCCCCAGAGAGCTGCTCTCTGAGGAGAG	2447
OS	QY	481	AGCCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATCTGAGTGGGA	540
XX	DB	2448	AGCCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATCTGAGTGGGA	2507
PN	QY	541	AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT	600
XX	DB	2508	AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT	2567
PI	QY	601	CCAGGAATCTCAAGAGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT	660
XX	DB	2568	CCAGGAATCTCAAGAGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT	2627
WPI	QY	661	CAAGGGATCTCGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAAGATACCTCGA	720
XX	DB	2628	CAAGGGATCTCGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAAGATACCTCGA	2687
DR	QY	721	GAAGAGTCAAGCACTTCCAGAGAGAAATGGCGCTCTGAAAGAGAACGTGAGCCACGTCAA	780
XX	DB	2688	GAAGAGTCAAGCACTTCCAGAGAGAAATGGCGCTCTGAAAGAGAACGTGAGCCACGTCAA	2747
XX	QY	781	TGACCTTCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACGGTATACCTCAGCAC	840
CC	DB	2748	TGACCTTCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACGGTATACCTCAGCAC	2807
CC	QY	841	TCTGGAAGACCTCAACACCCAGATGGAAGCTTCTCGAGGTGGCCGTGAGGACCCGAGTCAG	900
CC	DB	2808	TCTGGAAGACCTCAACACCCAGATGGAAGCTTCTCGAGGTGGCCGTGAGGACCCGAGTCAG	2867
CC	QY	901	CGAGCTGATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC	960
CC	DB	2868	CGAGCTGATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC	2927
CC	QY	961	TGTCAGGGTCCCTGGGAGAGAGCCATCTGCCCAAACAAGTGGCCCTACTATATCAACCA	1020
CC	DB	2928	TGTCAGGGTCCCTGGGAGAGAGCCATCTGCCCAAACAAGTGGCCCTACTATATCAACCA	2987
CC	QY	1021	CGAGACTCAAAACAATTGCTGGGACCATCCCAAATGACAGAGCTCTACAGTCTTTAGC	1080
CC	DB	2988	CGAGACTCAAAACAATTGCTGGGACCATCCCAAATGACAGAGCTCTACAGTCTTTAGC	3047
CC	QY	1081	TGACCTGAATATGTGATGATTTCTCAGCTTATAGGACTGCCATGAATCCGGAAGCTGCA	1140
CC	DB	3048	TGACCTGAATATGTGATGATTTCTCAGCTTATAGGACTGCCATGAATCCGGAAGCTGCA	3107
CC	QY	1141	GAAAGCCCTTTGCTTGGATCTCTTGAGCTGTGATGATGATGATGATGATGATGATGATGAT	1200
CC	DB	3108	GAAAGCCCTTTGCTTGGATCTCTTGAGCTGTGATGATGATGATGATGATGATGATGATGAT	3167
CC	QY	1201	CAACCTCAAGCAAAATGACAGCCCATGGATATCTCGCAGATTTAATTTCTTTGACCAC	1260
CC	DB	3168	CAACCTCAAGCAAAATGACAGCCCATGGATATCTCGCAGATTTAATTTCTTTGACCAC	3227
CC	QY	1261	TATTTATGACCGCTTGGAGCAGAGCAACAATTTGGTCAACGTCCTCTCTCGGTGGA	1320
CC	DB	3228	TATTTATGACCGCTTGGAGCAGAGCAACAATTTGGTCAACGTCCTCTCTCGGTGGA	3287

AD37234 standard; DNA; 3999 BP.

AD37234; (first entry)

21-AUG-2002

Human dystrophin minigene delta3990.

Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin; adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD; Becker muscular dystrophy; ds.

Homo sapiens.

WO200183695-A2.

08-NOV-2001.

27-APR-2001; 2001WO-US013677.

28-APR-2000; 2000US-0200777P.

(XIAO/) XIAO X.

Xiao X;

WPI; 2002-049342/06.

New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene.

Example 1; Page 46-47; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is human dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus, hinge H1 and rods R1 and R2), 8059-10227 (rods R2, R3 and R4, hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids)

Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 U; 0 Other;

Query Match 100.0%; Score 1501; DB 6; Length 3999;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ACAATGCAAGACTCCAAAGTGAATGAAGTCAACAGATGTTTATCAACCTGGA 60

1968 ACAATGCAAGACTCCAAAGTGAATGAAGTCAACAGATGTTTATCAACCTGGA 2027

61 TGAACAGCCCAAAATCTGAGATCCCTGGAGGTTCCGATGATGATGATGATGATGATGATGAT 120

2028 TGAACAGCCCAAAATCTGAGATCCCTGGAGGTTCCGATGATGATGATGATGATGATGATGAT 2087

121 AAGACGTTTGGATAACAATGAACTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACAT 180

2088 AAGACGTTTGGATAACAATGAACTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACAT 2147

181 TAGTCCCATTTGGAAGCCAGTTCTGACCAAGTGGAGGCTCTGCACTTTCTCTGAGCA 240

2148 TAGTCCCATTTGGAAGCCAGTTCTGACCAAGTGGAGGCTCTGCACTTTCTCTGAGCA 2207

QY 1321 TATGTGTCGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCGGTG 1380
 Db 3288 TATGTGTCGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCGGTG 3347
 QY 1381 CCGTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1440
 Db 3348 CCGTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 3407
 QY 1441 ATACCTTTTCAAGCAAGTGGCAGTTCACAGGATTTTGTGACCGGCGAGGCTGGGCT 1500
 Db 3408 ATACCTTTTCAAGCAAGTGGCAGTTCACAGGATTTTGTGACCGGCGAGGCTGGGCT 3467
 QY 1501 C 1501
 Db 3468 C 3468

RESULT 5
 AAD37230
 ID AAD37230 standard; DNA; 4182 BP.
 XX
 AC AAD37230;
 XX
 DT 21-AUG-2002 (first entry)
 XX
 Human dystrophin minigene delta4173.
 XX
 KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
 KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
 KW Becker muscular dystrophy; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200183695-A2.
 XX
 PD 08-NOV-2001.
 XX
 PF 27-APR-2001; 2001WO-US013677.
 XX
 PR 28-APR-2000; 2000US-0200777P.
 XX
 PA (XIAO/) XIAO X.
 XX
 PI Xiao X;
 XX
 DR WPI; 2002-049342/06.
 XX
 XX New dystrophin minigene for treating Duchenne or Becker muscular
 PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
 PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
 PT gene.
 XX
 PS Example 1; Page 43-44; 71pp; English.
 XX
 XX The present invention relates to an isolated nucleotide sequence encoding
 CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
 CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
 CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
 CC invention also relates to a recombinant adeno-associated virus (AAV)
 CC comprising dystrophin minigene operably linked to an expression control
 CC element. The dystrophin minigene in operable linkage with an expression
 CC control element, in a recombinant adeno-associated virus or retrovirus is
 CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
 CC dystrophy (BMD) in a mammalian subject. The present sequence is human
 CC dystrophin minigene delta4173 containing nucleotides 1-1592 (N-terminus,
 CC hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R2, R23 and R24,
 CC hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
 XX
 SQ Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1501; DB 6; Length 4182;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACATGCGACAGACCTCCAAAGTGAATTTGAAGCTCACACAGATGTTTATCACAACTGGA 60
 Db 2151 ACAATGCGACAGACCTCCAAAGTGAATTTGAAGCTCACACAGATGTTTATCACAACTGGA 2210
 QY 61 TGAACACAGACCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCGAGTCTCTGTACA 120
 Db 2211 TGAACACAGACCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCGAGTCTCTGTACA 2270
 QY 121 AAGACGTTTGGATACATGAACTTCAAGTGGAGTGAATTCGGAAGAAAGTCTCTCAACAT 180
 Db 2271 AAGACGTTTGGATACATGAACTTCAAGTGGAGTGAATTCGGAAGAAAGTCTCTCAACAT 2330
 QY 181 TAGTCCCATTTGGAAGCCAGTTCTGACCAAGTGGAGCGTCTGCACCTTTCTCTGCAGGA 240
 Db 2331 TAGTCCCATTTGGAAGCCAGTTCTGACCAAGTGGAGCGTCTGCACCTTTCTCTGCAGGA 2390
 QY 241 ACTTCTGCTGTGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGCACCTATTGGAGG 300
 Db 2391 ACTTCTGCTGTGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGCACCTATTGGAGG 2450
 QY 301 CGACTTTCCAGCAGTTCCAGAGCAGACGATGTACATAGGGCCCTTCAAGAGGGAATTGA 360
 Db 2451 CGACTTTCCAGCAGTTCCAGAGCAGACGATGTACATAGGGCCCTTCAAGAGGGAATTGA 2510
 QY 361 AACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACA 420
 Db 2511 AACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACA 2570
 QY 421 GCCTTTGGAAGGACTAGAGAACTCTACAGAGCCGAGAGCTGCTCTCTCAGGAGAG 480
 Db 2571 GCCTTTGGAAGGACTAGAGAACTCTACAGAGCCGAGAGCTGCTCTCTCAGGAGAG 2630
 QY 481 AGCCAGAGTGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 540
 Db 2631 AGCCAGAGTGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGA 2690
 QY 541 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 600
 Db 2691 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 2750
 QY 601 CCAGAACTTCAAGAGCCACGATGAGTGGACCTCAAGCTCGGCAAGCTGAGTGTAT 660
 Db 2751 CCAGAACTTCAAGAGCCACGATGAGTGGACCTCAAGCTCGGCAAGCTGAGTGTAT 2810
 QY 661 CAAGGATCTCGCAGCCCTGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCGA 720
 Db 2811 CAAGGATCTCGCAGCCCTGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCGA 2870
 QY 721 GAAAGTCAAGGCACTTCGAGGCAAAATTCGCTCTGAAAGAGAAAGCTGAGCCACCTCAA 780
 Db 2871 GAAAGTCAAGGCACTTCGAGGCAAAATTCGCTCTGAAAGAGAAAGCTGAGCCACCTCAA 2930
 QY 781 TGACCTTGTCTGCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAACCCTCAGCAC 840
 Db 2931 TGACCTTGTCTGCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAACCCTCAGCAC 2990
 QY 841 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTGAGGTGGCGTGGCGAGCCGAGTCA 900
 Db 2991 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTGAGGTGGCGTGGCGAGCCGAGTCA 3050
 QY 901 GCAGCTGATGAAGCCCAACAGGACCTTGGTCAGCATCTCAGCACTTTCTTTTCCACGTC 960
 Db 3051 GCAGCTGATGAAGCCCAACAGGACCTTGGTCAGCATCTCAGCACTTTCTTTTCCACGTC 3110
 QY 961 TGTCCAGGGTCTCTGGAGAGAGCCATCTCGCCAAACAAAGTGGCTCTATATACCA 1020
 Db 3111 TGTCCAGGGTCTCTGGAGAGAGCCATCTCGCCAAACAAAGTGGCTCTATATACCA 3170
 QY 1021 CGAGACTCAAAACAACTTGTGGGAGCCATCCAAAATGACAGAGCTCTACAGTCTTTAGC 1080
 Db 3171 CGAGACTCAAAACAACTTGTGGGAGCCATCCAAAATGACAGAGCTCTACAGTCTTTAGC 3230

301 CGACTTTCCAGAGTTCAGAGCAGACGATGTACATAGGGCTTCAAGAGGAATTGAA 360
Db |||||
2907 CGACTTTCCAGAGTTCAGAGCAGACGATGTACATAGGGCTTCAAGAGGAATTGAA 2966
QY |||||
361 AACTAAAGAACCTGTAACTATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGCA 420
Db |||||
2967 AACTAAAGAACCTGTAACTATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGCA 3026
QY |||||
421 GCCTTTGGAAGGACTAGAGAACTCTACAGAGAGCCAGAGACTGCCCTCTGAGGAGAG 480
Db |||||
3027 GCCTTTGGAAGGACTAGAGAACTCTACAGAGAGCCAGAGACTGCCCTCTGAGGAGAG 3086
QY |||||
481 AGCCCAAGATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 540
Db |||||
3087 AGCCCAAGATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 3146
QY |||||
541 AAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 600
Db |||||
3147 AAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 3206
QY |||||
601 CCAGGAACCTTCAAGAGGCCAGGATGAGTGGACCTCAAGCTCGGCCAAGCTGAGGTGAT 660
Db |||||
3207 CCAGGAACCTTCAAGAGGCCAGGATGAGTGGACCTCAAGCTCGGCCAAGCTGAGGTGAT 3266
QY |||||
661 CAAGGATCTGCGAGCCCGTGGCGGATCTCTCTATTGACTCTCTCAAGATCACTCGA 720
Db |||||
3267 CAAGGATCTGCGAGCCCGTGGCGGATCTCTCTATTGACTCTCTCAAGATCACTCGA 3326
QY |||||
721 GAAAGTCAGGCACTTCGAGGAGAAATTCGGCTCTGAAAGAGAGCTGAGCCAGCTCAA 780
Db |||||
3327 GAAAGTCAGGCACTTCGAGGAGAAATTCGGCTCTGAAAGAGAGCTGAGCCAGCTCAA 3386
QY |||||
781 TGACCTTGCTCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAAAGCTCAGCAC 840
Db |||||
3387 TGACCTTGCTCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAAAGCTCAGCAC 3446
QY |||||
841 TCTGGAAGACCTGAAACACAGATGAGAGCTTTCGAGGTGGCGTGGAGGACCGAGTCAG 900
Db |||||
3447 TCTGGAAGACCTGAAACACAGATGAGAGCTTTCGAGGTGGCGTGGAGGACCGAGTCAG 3506
QY |||||
901 GCAGCTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTC 960
Db |||||
3507 GCAGCTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTC 3566
QY |||||
961 TGTCCAGGTTCCTGGAGAGGCCATCTCGCCAAACAAAGTGCCCTACTATCAACA 1020
Db |||||
3567 TGTCCAGGTTCCTGGAGAGGCCATCTCGCCAAACAAAGTGCCCTACTATCAACA 3626
QY |||||
1021 CGAGACTCAACAACTTGTGGGACCATCCCAAAATGACAGACTCTACAGTCTTTAGC 1080
Db |||||
3627 CGAGACTCAACAACTTGTGGGACCATCCCAAAATGACAGACTCTACAGTCTTTAGC 3686
QY |||||
1081 TGACCTGAATAATGTCAAGTTCAGCTTATAGGACTGCCATGAACCTCCGAAGACTGCA 1140
Db |||||
3687 TGACCTGAATAATGTCAAGTTCAGCTTATAGGACTGCCATGAACCTCCGAAGACTGCA 3746
QY |||||
1141 GAAGCCCTTTCTTGGATCTCTGAGCTGTGACGTGCAATGATGCTTGGACCGAGCA 1200
Db |||||
3747 GAAGCCCTTTCTTGGATCTCTGAGCTGTGACGTGCAATGATGCTTGGACCGAGCA 3806
QY |||||
1201 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCTCGAGATTATTAATGTTTGGCCAC 1260
Db |||||
3807 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCTCGAGATTATTAATGTTTGGCCAC 3866
QY |||||
1261 TATTTATGACCCCTGGAGAGAGCAACAATTTGGTCAACGTCCTCTCGCTGGA 1320
Db |||||
3867 TATTTATGACCCCTGGAGAGAGCAACAATTTGGTCAACGTCCTCTCGCTGGA 3926
QY |||||
1321 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGAGCAACAGGAGGATCCGTGT 1380
Db |||||
3927 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGAGCAACAGGAGGATCCGTGT 3986
QY |||||
1381 CCTGTCTTTTAAACTGGCATATTTCCCTGTGTAAAGCACATTTTGAAGAGCAAGTACAG 1440

Db |||||
3987 CCTGTCTTTTAAACTGGCATATTTCCCTGTGTAAAGCACATTTGGAAGACAGTACAG 4046
QY |||||
1441 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACCGCGCAGGCTGGGCT 1500
Db |||||
4047 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACCGCGCAGGCTGGGCT 4106
QY |||||
1501 C 1501
Db |||||
4107 C 4107

RESULT 9
AAD37256
ID AAD37256 standard; DNA; 4966 BP.
XX
AC AAD37256;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta3990.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
PN W0200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 59-60; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 U; 0 Other;
Query Match 100.0%; Score 1501; DB 6; Length 4966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAATGGCAAGACCTTCCAAGGTGAAATTTGAAGCTCACAGATGTTTATCAACCTGGA 60
Db 2725 ACAATGGCAAGACCTTCCAAGGTGAAATTTGAAGCTCACAGATGTTTATCAACCTGGA 2784

[illegible]

CC invention also relates to a recombinant adeno-associated virus (AAV)
 CC comprising dystrophin minigene operably linked to an expression control
 CC element. The dystrophin minigene in operable linkage with an expression
 CC control element, in a recombinant adeno-associated virus or retrovirus is
 CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
 CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
 CC vector plasmid construct containing human dystrophin minigenes, a
 CC cytomegalovirus (CMV) promoter and a small polyA signal sequence
 XX
 SQ Sequence 4990 BP; 1439 A; 1185 C; 1208 G; 1158 T; 0 U; 0 Other;

Query Match 100.0%; Score 1501; DB 6; Length 4990;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAATGGCAAGACTCCCAAGTGAATTTGAAGTCTCACAGATGTTTATCAACAACCTGGA 60
 DB 2749 ACAATGGCAAGACTCCCAAGTGAATTTGAAGTCTCACAGATGTTTATCAACAACCTGGA 2808

QY 61 TGAATAACAGCCAAATAATCTGAGATCCCTGGAGGTTCCGATGATGCTCTGTTCACA 120
 DB 2809 TGAATAACAGCCAAATAATCTGAGATCCCTGGAGGTTCCGATGATGCTCTGTTCACA 2868

QY 121 AAGACGTTTGGATAAATGATGAATTCAGTGGAGTGAATTCGGAAGAAATCTCTCAACAT 180
 DB 2869 AAGACGTTTGGATAAATGATGAATTCAGTGGAGTGAATTCGGAAGAAATCTCTCAACAT 2928

QY 181 TAGTCCCATTTGGAAGCCAGTCTTGACAGCTGGAAGCGTCTGCACCTTTCTCTGCAGGA 240
 DB 2929 TAGTCCCATTTGGAAGCCAGTCTTGACAGCTGGAAGCGTCTGCACCTTTCTCTGCAGGA 2988

QY 241 ACTTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCACCTTATTTGGAGG 300
 DB 2989 ACTTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCACCTTATTTGGAGG 3048

QY 301 CGACTTTCCAGCAGTTCAGAAAGCAGAAACGATGTACATAGGGCCCTTCAAGAGGGAATTCGA 360
 DB 3049 CGACTTTCCAGCAGTTCAGAAAGCAGAAACGATGTACATAGGGCCCTTCAAGAGGGAATTCGA 3108

QY 361 AACTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTGCAATATTTCTGACAGAGCA 420
 DB 3109 AACTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTGCAATATTTCTGACAGAGCA 3168

QY 421 GCCTTTGGAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCTCTGAGGAGAG 480
 DB 3169 GCCTTTGGAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCTCTGAGGAGAG 3228

QY 481 AGCCAGAAATGTCACTGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACCTGAGTGGGA 540
 DB 3229 AGCCAGAAATGTCACTGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACCTGAGTGGGA 3288

QY 541 AAAATTTGAACCTGCACTCGCTGACTGCGCAGAGAAATAGATGAGACCTTTGAAGACT 600
 DB 3289 AAAATTTGAACCTGCACTCGCTGACTGCGCAGAGAAATAGATGAGACCTTTGAAGACT 3348

QY 601 CCAGGAACCTTCAAGAGGCCACGATGAGCTGAGCTCAAGCTGCGCCCAAGCTGAGGTGAT 660
 DB 3349 CCAGGAACCTTCAAGAGGCCACGATGAGCTGAGCTCAAGCTGCGCCCAAGCTGAGGTGAT 3408

QY 661 CAGAGGATCCTGGCAGCCGCTGGGGGATCTCTCTATTGACTCTCTCCAAAGATCACTCGA 720
 DB 3409 CAGAGGATCCTGGCAGCCGCTGGGGGATCTCTCTATTGACTCTCTCCAAAGATCACTCGA 3468

QY 721 GAAAGTCAAGGCACATTCAGAGGAAATTCGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 780
 DB 3469 GAAAGTCAAGGCACATTCAGAGGAAATTCGCGCTCTGAAAGAGAACGTGAGCCACGTCAA 3528

QY 781 TGACCTTGTCTGCCAGCTTACACTTTTGGGCAATTCAGCTCTCAACCGTATTAACCTCAGCAC 840
 DB 3529 TGACCTTGTCTGCCAGCTTACACTTTTGGGCAATTCAGCTCTCAACCGTATTAACCTCAGCAC 3588

QY 841 TCTGGAAGACCTGAACACAGATGGAAGCTTCTCAGGTGCGCTCGAGGACCGAGTTCAG 900

DB 3589 TCTGGAAGACCTGAACACACCATGGAAGCTTCTGCAAGTGGCCCTCGAGGACCGAGTTCAG 3648
 QY 901 GCAGCTGATGAAGCCACACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTC 960
 DB 3649 GCAGCTGATGAAGCCACACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTC 3708
 QY 961 TGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 1020
 DB 3709 TGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 3768
 QY 1021 CGAGACTCAAAACACTTCTGCTGGGACCATCCCAAAATGACAGAGCTTACCAAGCTCTTAGC 1080
 DB 3769 CGAGACTCAAAACACTTCTGCTGGGACCATCCCAAAATGACAGAGCTTACCAAGCTCTTAGC 3828
 QY 1081 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA 1140
 DB 3829 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA 3888
 QY 1141 GAAGGCCCTTTGCTTGGATCTCTTGACCTGTGAGCTGTCATGTGATGCCCTTGGACACGA 1200
 DB 3889 GAAGGCCCTTTGCTTGGATCTCTTGACCTGTGAGCTGTCATGTGATGCCCTTGGACACGA 3948
 QY 1201 CAACCTCAAGAAATAATGACAGGCCATGATATCTCTGACAGATTATTAATTTGTTGACCAC 1260
 DB 3949 CAACCTCAAGAAATAATGACAGGCCATGATATCTCTGACAGATTATTAATTTGTTGACCAC 4008
 QY 1261 TATTTATGACCGCTGGAGCAAGACACAAATTTGGTCAACGCTCCCTCTCTGCGTGA 1320
 DB 4009 TATTTATGACCGCTGGAGCAAGACACAAATTTGGTCAACGCTCCCTCTCTGCGTGA 4068
 QY 1321 TATGTGCTCAACTGGCTGCTGAATTTATGATACGGGACGAAACAGGAGGATCCGTGT 1380
 DB 4069 TATGTGCTCAACTGGCTGCTGAATTTATGATACGGGACGAAACAGGAGGATCCGTGT 4128
 QY 1381 CCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAGCACATTTTGGAGACAGTACAG 1440
 DB 4129 CCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAGCACATTTTGGAGACAGTACAG 4188
 QY 1441 ATACCTTTTCAAGCAAGTGGCAGTTCACAGAGATTGTGACCCAGCGCAGGCTGGGCT 1500
 DB 4189 ATACCTTTTCAAGCAAGTGGCAGTTCACAGAGATTGTGACCCAGCGCAGGCTGGGCT 4248
 QY 1501 C 1501
 DB 4249 C 4249

RESULT 11

AAD37264

ID AAD37264 standard; DNA; 5060 BP.

XX AAD37264;

XX AC

XX XX

DT 21-AUG-2002 (first entry)

XX

XX Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.

DE

XX

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX

XX Becker muscular dystrophy; ds.

OS

OS Homo sapiens.

OS Cytomegalovirus.

OS Unidentified.

OS Chimeric.

XX

PN W0200183695-A2.

XX

XX 08-NOV-2001.

XX

PF 27-APR-2001; 2001WO-US013677.

XX

PR 28-APR-2000; 2000US-0200777P.

XX Adeno-associated virus vector plasmid, AAV-MCK-delta173.
 DE
 XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
 KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
 KW Becker muscular dystrophy; ds.
 XX
 OS Homo sapiens.
 OS Unidentified.
 OS Chimeric.
 XX
 XX W0200183695-A2.
 XX
 XX 08-NOV-2001.
 XX
 XX 27-APR-2001; 2001WO-US013677.
 XX
 XX 28-APR-2000; 2000US-0200777P.
 XX
 XX (XIAO/) XIAO X.
 XX
 XX Xiao X;
 PI
 XX WPI; 2002-049342/06.
 XX
 XX New dystrophin minigene for treating Duchenne or Becker muscular
 PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
 PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
 PT gene.
 XX
 XX Example 1; Page 57-59; 71pp; English.
 XX
 XX The present invention relates to an isolated nucleotide sequence encoding
 CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
 CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
 CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
 CC invention also relates to a recombinant adeno-associated virus (AAV)
 CC comprising dystrophin minigene operably linked to an expression control
 CC element. The dystrophin minigene in operable linkage with an expression
 CC control element, in a recombinant adeno-associated virus or retrovirus is
 CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
 CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
 CC vector plasmid construct containing human dystrophin minigenes, a muscle
 CC creatine kinase (CK) promoter and a small polyA signal sequence
 XX
 XX Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 1501; DB 6; Length 5149;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACAATGGCAGACCTCCAGGTTGAATTCAGCTCACACAGATGTTTATCAACCTGGA 60
 DB 2908 ACAATGGCAGACCTCCAGGTTGAATTCAGCTCACACAGATGTTTATCAACCTGGA 2967
 QY 61 TGAACAACGCCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGCGTCTCTGTACA 120
 DB 2968 TGAACAACGCCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGCGTCTCTGTACA 3027
 QY 121 AAGAGGTTTGGATACATGAACTTCAAGTGGAGTGAATTCGGAAGAAAGTCTCTCAACAT 180
 DB 3028 AAGAGGTTTGGATACATGAACTTCAAGTGGAGTGAATTCGGAAGAAAGTCTCTCAACAT 3087
 QY 181 TAGGTCCCATTTGGAAGCCAGTTCTGACAGTGAAGCGTCTGCACCTTTCTCTGCAGGA 240
 DB 3088 TAGGTCCCATTTGGAAGCCAGTTCTGACAGTGAAGCGTCTGCACCTTTCTCTGCAGGA 3147
 QY 241 ACTTCTGGTGTGGCTTACAGCTGGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG 300
 DB 3148 ACTTCTGGTGTGGCTTACAGCTGGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG 3207
 QY 301 CGACTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGCGCTTCAGAGGGAATTGAA 360

DB 3208 CGACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGCGCTTCAGAGGGAATTGAA 3267
 QY 361 AACTAAGAAGCCTGTAATCATATGAGTACTCTTTGAGACTGTGCGAATATTTCTGACAGAGCA 420
 DB 3268 AACTAAGAAGCCTGTAATCATATGAGTACTCTTTGAGACTGTGCGAATATTTCTGACAGAGCA 3327
 QY 421 GCCTTTTGAAGGACTTAGAGAACTCTACAGAGGCCAGAGAGCTGCTCTCTGAGGAGAG 480
 DB 3328 GCCTTTTGAAGGACTTAGAGAACTCTACAGAGGCCAGAGAGCTGCTCTCTGAGGAGAG 3387
 QY 481 AGCCCAAGATGTCACTCGGCTTTACGAAAGCAGGCTGAGAGGTCATACTAGTGGGA 540
 DB 3388 AGCCCAAGATGTCACTCGGCTTTACGAAAGCAGGCTGAGAGGTCATACTAGTGGGA 3447
 QY 541 AAAATTGAACCTGCATCCGCTGACTGGCAGAGAAATAGATGAGACCTTTGAAGACT 600
 DB 3448 AAAATTGAACCTGCATCCGCTGACTGGCAGAGAAATAGATGAGACCTTTGAAGACT 3507
 QY 601 CCAGGAACCTTCAAGAGGCCACGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 660
 DB 3508 CCAGGAACCTTCAAGAGGCCACGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 3567
 QY 661 CAAGGGATCTGGCAGGCCCTGGCGGATCTCTCATGACTCTCTCAAGATCACCTCGA 720
 DB 3568 CAAGGGATCTGGCAGGCCCTGGCGGATCTCTCATGACTCTCTCAAGATCACCTCGA 3627
 QY 721 GAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCTCTGAAAGAGAAAGTGAAGCACCTCAA 780
 DB 3628 GAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCTCTGAAAGAGAAAGTGAAGCACCTCAA 3687
 QY 781 TGACCTTGTCTGGCAGCTTACCACTTTGGGCAATCAGCTCTCAACCATAAACCCTCAGCAC 840
 DB 3688 TGACCTTGTCTGGCAGCTTACCACTTTGGGCAATCAGCTCTCAACCATAAACCCTCAGCAC 3747
 QY 841 TCTGGAAGACCTCAACACAGATGGAAGCTTCTGCAAGTGGCGCTGAGAGCCGAGTCA 900
 DB 3748 TCTGGAAGACCTCAACACAGATGGAAGCTTCTGCAAGTGGCGCTGAGAGCCGAGTCA 3807
 QY 901 CGAGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 960
 DB 3808 CGAGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 3867
 QY 961 TGTCCAGGTTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTCCCTACTATATCAACA 1020
 DB 3868 TGTCCAGGTTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTCCCTACTATATCAACA 3927
 QY 1021 CGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGCTCTTTAGC 1080
 DB 3928 CGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGCTCTTTAGC 3987
 QY 1081 TGACCTGAATATGTCAGATTTCTCAGCTTATAGGACTGCCATGAACTCCGAAAGACTGA 1140
 DB 3988 TGACCTGAATATGTCAGATTTCTCAGCTTATAGGACTGCCATGAACTCCGAAAGACTGA 4047
 QY 1141 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGTGTGATGCTTGGACACGCA 1200
 DB 4048 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGTGTGATGCTTGGACACGCA 4107
 QY 1201 GAACCTCAAGCAAAATGACAGCCCATGATATCTCGAGATTTAATATTTGTTGACCAC 1260
 DB 4108 GAACCTCAAGCAAAATGACAGCCCATGATATCTCGAGATTTAATATTTGTTGACCAC 4167
 QY 1261 TATTTATGACCCCTTGGAGCAGAGCAACAAATTTGTTCAAGCTCCCTCTCTCGGTGA 1320
 DB 4168 TATTTATGACCCCTTGGAGCAGAGCAACAAATTTGTTCAAGCTCCCTCTCTCGGTGA 4227
 QY 1321 TATGTCCTGAACCTGGCTGCTGAATTTATATGATACGGGACGAAACAGGAGGATCCGTTG 1380
 DB 4228 TATGTCCTGAACCTGGCTGCTGAATTTATATGATACGGGACGAAACAGGAGGATCCGTTG 4287
 QY 1381 CCTGTCTTTTAAACCTGGCATCTTCCCTGTGTAAGACATTTGGAGACAGATACAG 1440
 DB 4288 CCTGTCTTTTAAACCTGGCATCTTCCCTGTGTAAGACATTTGGAGACAGATACAG 4347

QY 1441 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGAGGCTGGGCT 1500
DB 4348 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGAGGCTGGGCT 4407
QY 1501 C 1501
DB 4408 C 4408
RESULT 13
ID ABK81998 standard; DNA; 5339 BP.
XX AC ABK81998;
XX DT 13-AUG-2002 (first entry)
XX DE DNA encoding mini-dystrophin protein deltaR2-R21.
XX KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
XX KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200229056-A2.
XX XX WO200229056-A2.
XX PD 11-APR-2002.
XX PF 04-OCT-2001; 2001WO-US031126.
XX PR 06-OCT-2000; 2000US-0238848P.
XX XX (UNMI) UNIV MICHIGAN.
XX PA Chamberlain JS, Harper SQ;
XX PI WPI; 2002-435334/46.
XX DR A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains, or
PT a nucleic acid sequence encoding the mini-dystrophin peptide.
XX Example 6; Fig 13; 145pp; English.
XX The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n spectrin-
CC like repeats, where n is an even number between 4-24, or a nucleic acid
CC encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC polynucleotide encoding it is useful as a medicament, for preparing a
CC drug for therapeutic application and in the preparation of a composition
CC for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC (DMD). This sequence represents a mini-dystrophin sequence of the
XX invention
SQ Sequence 5339 BP; 1638 A; 1191 C; 1187 G; 1323 T; 0 U; 0 Other;
Query Match 100.0%; Score 1501; DB 6; Length 5339;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAATGCAAGACCTCCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCACAACTGGA 60
DB 1693 ACAATGCAAGACCTCCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCACAACTGGA 1752
QY 61 TGAAGACAGCCAAATCTGAGATCCCTGGAGGTCCTGATGCGAGTCTGTGTACA 120
DB 1753 TGAAGACAGCCAAATCTGAGATCCCTGGAGGTCCTGATGCGAGTCTGTGTACA 1812
QY 121 AAGACGTTTGGTAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAGAACTCTCTCAAACAT 180

DB 1813 AAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAAACAT 1872
QY 181 TAGTCTCCATTTTGAAGCCAGTTCTGACCAAGTGGAGCGCTCTGCACCTTTCTCTGCAGGA 240
DB 1873 TAGGTCCCATTTTGAAGCCAGTTCTGACCAAGTGGAGCGCTCTGCACCTTTCTCTGCAGGA 1932
QY 241 ACTTCTGGTGTGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGGCACTTATTGGAGG 300
DB 1933 ACTTCTGGTGTGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGGCACTTATTGGAGG 1992
QY 301 CGACTTTTCCAGCAGTTTCAAGACAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAA 360
DB 1993 CGACTTTTCCAGCAGTTTCAAGACAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAA 2052
QY 361 AACTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 420
DB 2053 AACTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 2112
QY 421 GCCTTTTGGAGGAGTACAGAACTCTACAGAGAGCCAGAGAGCTGCCTTCCTGAGGAGAG 480
DB 2113 GCCTTTTGGAGGAGTACAGAACTCTACAGAGAGCCAGAGAGCTGCCTTCCTGAGGAGAG 2172
QY 481 AGCCGAGATGTCACTCGGCTTCTAGAAAAGCAGGCTGAGGAGGTCATATCTGAGTGGGA 540
DB 2173 AGCCGAGATGTCACTCGGCTTCTAGAAAAGCAGGCTGAGGAGGTCATATCTGAGTGGGA 2232
QY 541 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGACACCTTTGAAAAGACT 600
DB 2233 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGACACCTTTGAAAAGACT 2292
QY 601 CAGGAACTTCAAGAGGCAACGGATGAGCTGAGACTCAAGCTGCGCAGGCTGAGGTTGAT 660
DB 2293 CAGGAACTTCAAGAGGCAACGGATGAGCTGAGACTCAAGCTGCGCAGGCTGAGGTTGAT 2352
QY 661 CAAGGGATCTGCGCAGCGCTGGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCGA 720
DB 2353 CAAGGGATCTGCGCAGCGCTGGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCGA 2412
QY 721 GAAAGTCAAGCACTTTCGAGGAGAAATTTGGCTCTGAAAAGAGACGTTGAGGCCACGTCOA 780
DB 2413 GAAAGTCAAGCACTTTCGAGGAGAAATTTGGCTCTGAAAAGAGACGTTGAGGCCACGTCOA 2472
QY 781 TGACCTTGTCTGCGCAGCTTACCACTTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 840
DB 2473 TGACCTTGTCTGCGCAGCTTACCACTTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCAC 2532
QY 841 TCTGGAAGACCTGAAACACAGATGGAAGTTCTGAGGTGGCGCTCGAGGACCGAGTCAG 900
DB 2533 TCTGGAAGACCTGAAACACAGATGGAAGTTCTGAGGTGGCGCTCGAGGACCGAGTCAG 2592
QY 901 GCAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC 960
DB 2593 GCAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC 2652
QY 961 TGTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 1020
DB 2653 TGTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 2712
QY 1021 CGAGACTCAAAACAACTTGTGGAGCCATCCCAAATGACAGAGCTTACCAGTCTTTAGC 1080
DB 2713 CGAGACTCAAAACAACTTGTGGAGCCATCCCAAATGACAGAGCTTACCAGTCTTTAGC 2772
QY 1081 TGACCTGAATTAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 2773 TGACCTGAATTAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2832
QY 1141 GAAAGCCCTTTGGTCTGATCTCTTGGAGCTGTGAGCTGATGATGATGATGATGATGATGATGAT 1200
DB 2833 GAAAGCCCTTTGGTCTGATCTCTTGGAGCTGTGAGCTGATGATGATGATGATGATGATGATGAT 2892
QY 1201 CAACCTCAAGCAAAATGACAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 2893 CAACCTCAAGCAAAATGACAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2952

QY 1261 TATTATGACCGCTGAGCAAGACGACAAATTTGGTCAACGTCCTCTCTGGGTGGA 1320
 DB 2953 TATTATGACCGCTGAGCAAGACGACAAATTTGGTCAACGTCCTCTCTGGGTGGA 3012
 QY 1321 TATGTGCTGAACCTGGCTGCTGAATGTTATGATACGGGACGACAGGAGGATCCGTGT 1380
 DB 3013 TATGTGCTGAACCTGGCTGCTGAATGTTATGATACGGGACGACAGGAGGATCCGTGT 3072
 QY 1391 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGGAGACAGATACAG 1440
 DB 3073 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGGAGACAGATACAG 3132
 QY 1441 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGGAGGCTGGGCT 1500
 DB 3133 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGGAGGCTGGGCT 3192
 QY 1501 C 1501
 DB 3193 C 3193

RESULT 14

ABK81999
 ID ABK81999 standard; DNA; 5462 BP.
 XX AC ABK81999;
 XX DT 13-AUG-2002 (first entry)
 XX DE DNA encoding mini-dystrophin protein deltaR2-R21+H3.
 XX KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
 XX KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
 XX OS Homo sapiens.
 XX OS Synthetic.
 XX FN WO200229056-A2.
 XX PD 11-APR-2002.
 XX PF 04-OCT-2001; 2001WO-US031126.
 XX PR 06-OCT-2000; 2000US-0238848P.
 XX PA (UNMI) UNIV MICHIGAN.
 XX PI Chamberlain JS, Harper SQ;
 XX DR WPI; 2002-435334/46.
 XX PT A composition for preparing therapeutic drugs, has a mini-dystrophin
 PT peptide comprising a specific number of spectrin-like repeat domains, or
 PT a nucleic acid sequence encoding the mini-dystrophin peptide.
 XX PS Disclosure; Fig 14; 145pp; English.
 XX CC The invention describes a composition comprising a mini-dystrophin
 CC peptide comprising a spectrin-like repeat domain, where the domain
 CC comprises n spectrin-like repeats, and contains no more than n spectrin-
 CC like repeats, where n is an even number between 4-24, or a nucleic acid
 CC encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
 CC polynucleotide encoding it is useful as a medicament, for preparing a
 CC drug for therapeutic application and in the preparation of a composition
 CC for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
 CC (DMD). This sequence represents a mini-dystrophin sequence of the
 CC invention
 XX SQ Sequence 5462 BP; 1668 A; 1225 C; 1212 G; 1357 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1501; DB 6; Length 5462;
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACAATGGCAAGACCTCCAAAGGTGAATTTGAAGCTCACACAGATGTTTATCAACAACCTGGA 60
 DB 1816 ACATGGCAGACCTCCAAAGGTGAATTTGAAGCTCACACAGATGTTTATCAACAACCTGGA 1875
 QY 61 TGAACACAGCCAAATAATCCTCAGATCCCTGGAAGTTCGATGATGATGATGATGATGATGAT 120
 DB 1876 TGAACACAGCCAAATAATCCTCAGATCCCTGGAAGTTCGATGATGATGATGATGATGATGAT 1935
 QY 121 AAGACGTTTGGATTAACATGAACCTTCAAGTGGAGTGAATTCGGAAGAAAGTCTCTCAACAT 180
 DB 1936 AAGACGTTTGGATTAACATGAACCTTCAAGTGGAGTGAATTCGGAAGAAAGTCTCTCAACAT 1995
 QY 181 TAGGTCCCATTTTGGAGGACGATTCGACGAGTGGAGGCTCTGACACCTTTCTCTGAGGA 240
 DB 1996 TAGGTCCCATTTTGGAGGACGATTCGACGAGTGGAGGCTCTGACACCTTTCTCTGAGGA 2055
 QY 241 ACTTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG 300
 DB 2056 ACTTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG 2115
 QY 301 CGACTTTCAGCAGTTCAGACGACGATGTACATAGGGCTTCAAGAGGGGAATTGAA 360
 DB 2116 CGACTTTCAGCAGTTCAGACGACGATGTACATAGGGCTTCAAGAGGGGAATTGAA 2175
 QY 361 AACTAAGAACCTGTAAATCATCAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 420
 DB 2176 AACTAAGAACCTGTAAATCATCAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 2235
 QY 421 GCCTTTGGAAGCACTAGAGAACTCTACAGAGCCGACAGAGCTGCTCTCTGAGGAGAG 480
 DB 2236 GCCTTTGGAAGCACTAGAGAACTCTACAGAGCCGACAGAGCTGCTCTCTGAGGAGAG 2295
 QY 481 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTCAGGAGGTCAATATCTAGTGGGA 540
 DB 2296 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTCAGGAGGTCAATATCTAGTGGGA 2355
 QY 541 AAAATTGAACCTGCACTCGGCTGCTGAGAGAAATAATAGATGAGACCTTTGAAAGACT 600
 DB 2356 AAAATTGAACCTGCACTCGGCTGCTGAGAGAAATAATAGATGAGACCTTTGAAAGACT 2415
 QY 601 CAGGAACTTCAAGAGCCGACGATGAGCTGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 660
 DB 2416 CAGGAACTTCAAGAGCCGACGATGAGCTGACCTCAAGCTGCGCCCAAGCTGAGGTGAT 2475
 QY 661 CAAGGATCTGCGACGCTGGGCGATCTCTCATTCATCTCTCTCAAGATCACTCGA 720
 DB 2476 CAAGGATCTGCGACGCTGGGCGATCTCTCATTCATCTCTCTCAAGATCACTCGA 2535
 QY 721 GAAAGTCAAGGCACTTCGAGGAGAAATTCGGCTCTGAAAGAGACGCTGAGCCAGTCAA 780
 DB 2536 GAAAGTCAAGGCACTTCGAGGAGAAATTCGGCTCTGAAAGAGACGCTGAGCCAGTCAA 2595
 QY 781 TGACCTTGTCTGCGACGCTTACCACTTTGGGCAATTCAGCTCTCAAGTATACCTCAGCAC 840
 DB 2596 TGACCTTGTCTGCGACGCTTACCACTTTGGGCAATTCAGCTCTCAAGTATACCTCAGCAC 2655
 QY 841 TCTGGAGACCTTGAACACGATGGAAGTCTCTGAGGTGGCGCTGAGGAGCCAGTCAAG 900
 DB 2656 TCTGGAGACCTTGAACACGATGGAAGTCTCTGAGGTGGCGCTGAGGAGCCAGTCAAG 2715
 QY 901 GCAGCTGCATGAAGCCGACGAGGATTTGGTTCAGCATCTCAGCACTTTCTTCTCCAGTC 960
 DB 2716 GCAGCTGCATGAAGCCGACGAGGATTTGGTTCAGCATCTCAGCACTTTCTTCTCCAGTC 2775
 QY 961 TGTCCAGGTCCTCTGGAGAGAGCCATCTCGGCAAAAGTGGCCTACTATATCAACCA 1020
 DB 2776 TGTCCAGGTCCTCTGGAGAGAGCCATCTCGGCAAAAGTGGCCTACTATATCAACCA 2835
 QY 1021 CGAGACTCAAAACAACTTGTGTGGGACCATCCCAAAATGACAGAGCTCTACAGCTCTTAGC 1080
 DB 2836 CGAGACTCAAAACAACTTGTGTGGGACCATCCCAAAATGACAGAGCTCTACAGCTCTTAGC 2895

QY	1081	TGACCTGATATATGTCAGATTCTCAGCTTATAGACTGCGCATGAACCTCCGAAGACTGCA	1140
Db	2896	TGACCTGATATATGTCAGATTCTCAGCTTATAGACTGCGCATGAACCTCCGAAGACTGCA	2955
QY	1141	GAAGGCCCTTTTGGTTCCTTTGAGCTGTCTTTGAGCCCTGTGATGCTGATGCTTGGACCA	1200
Db	2956	GAAGGCCCTTTTGGTTCCTTTGAGCCCTGTGATGCTGATGCTTGGACCA	3015
QY	1201	CAACCTCAGCAAAATGACCAAGCCATGCGATATCTTCGAGATATTAATTTGTTGACCA	1260
Db	3016	CAACCTCAGCAAAATGACCAAGCCATGCGATATCTTCGAGATATTAATTTGTTGACCA	3075
QY	1261	TATTATGACCGCTTGGAGCAAGACGACAAACAATTTGGTCAACGTCCTCTCTGCGTGA	1320
Db	3076	TATTATGACCGCTTGGAGCAAGACGACAAACAATTTGGTCAACGTCCTCTCTGCGTGA	3135
QY	1321	TATGTGCTGAGCAAGTGGCTGTGATGTTTATGATACGGACGACAGGAGGATCCGCTGT	1380
Db	3136	TATGTGCTGAGCAAGTGGCTGTGATGTTTATGATACGGACGACAGGAGGATCCGCTGT	3195
QY	1381	CTGTCTCTTTAAACTGGCATCATTTCCCTGTGTTAAAGCACATTTGGAAGCAAGTACAG	1440
Db	3196	CTGTCTCTTTAAACTGGCATCATTTCCCTGTGTTAAAGCACATTTGGAAGCAAGTACAG	3255
QY	1441	ATACCTTTTCAGCAAGTGGCAAGTTCACACAGGATTTTGTGACGACGCGCGCTGGCGCT	1500
Db	3256	ATACCTTTTCAGCAAGTGGCAAGTTCACACAGGATTTTGTGACGACGCGCGCTGGCGCT	3315
QY	1501	C 1501	
Db	3316	C 3316	
RESULT 15			
AAD06794			
ID	AAD06794 standard; DNA; 5952 BP.		
XX	AAD06794;		
XX	06-AUG-2001 (first entry)		
XX	Human dystrophin gene (Becker form).		
XX	Human; dystrophin; extein; intein; trans-splicing; gene therapy;		
KW	Duchenne muscular dystrophy; Becker muscular dystrophy; DMD; BMD; ds.		
XX	Homo sapiens.		
OS	Location/Qualifiers		
Key	2847..2848		
PH	misc_recomb		
FT	/tag= a		
FT	/label= S4_junction_site		
FT	/note= "Dystrophin gene is split at this site and attached to an intein sequence. The resulting fragment is used to produce plasmid pSD4"		
FT	2952..2953		
FT	misc_recomb		
PT	/tag= b		
PT	/label= S3_junction_site		
PT	/note= "Dystrophin gene is split at this site and attached to an intein sequence. The resulting fragment is used to produce plasmid pSD3"		
FT	3198..3199		
FT	misc_recomb		
PT	/tag= c		
PT	/label= S2_junction_site		
FT	/note= "Dystrophin gene is split at this site and attached to an intein sequence. The resulting fragment is used to produce plasmid pSD2"		
FT	3300..3301		
FT	misc_recomb		
PT	/tag= d		
PT	/label= S1_junction_site		
FT	/note= "Dystrophin gene is split at this site and attached to an intein sequence. The resulting fragment is used to produce plasmid pSD1"		

PT		used to produce plasmid pSD1"
XX		
PN	W0200129243-A1.	
XX		
PD	26-APR-2001.	
XX		
PF	13-OCT-2000; 2000WC-CA001216.	
XX		
PR	15-OCT-1999; 99US-0159868P.	
XX		
PA	(UYDA-) UNIV DALHOUSIE.	
FA	(UYPI-) UNIV PITTSBURGH.	
XX		
PI	Paul XL, Xiao X;	
XX		
DR	WPI; 2001-367297/38.	
XX		
PT	Use of spontaneous or automatic protein splicing to join two or more peptides at junction site involves expressing extein peptides having co-	
PT	reacting portions of split intein attached to them, so that peptides splice.	
XX		
PS	Example 1; Fig 2; 8lpp; English.	
CC	The invention relates to a method directed to the use of spontaneous or	
CC	automatic protein trans-splicing to join two or more peptides at junction	
CC	site. This method involves expressing extein peptides having co-reacting	
CC	portions of split intein attached to them, so that peptides will splice	
CC	automatically under suitable conditions. The invention also provides	
CC	methods for circumventing virion packaging size limitations in	
CC	recombinant virus particle, by splitting a coding region for a protein to	
CC	be delivered into two or more extein genes, which are packaged in	
CC	separate virus particles and are co-delivered in a target cell for the	
CC	expression and for subsequent trans-splicing to form the complete	
CC	protein. In particular, the method is used for trans-splicing human	
CC	dystrophin and in gene therapies of recombinant adeno-associated virus	
CC	(AAV) particles that encode trans-spliced dystrophin, for treating	
CC	diseases such as Duchenne muscular dystrophy (DMD) or Becker muscular	
CC	dystrophy (BMD). The present sequence is human dystrophin gene (Becker	
CC	form)	
XX		
SQ	Sequence 5952 BP; 1860 A; 1344 C; 1410 G; 1338 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 1501; DB 5; Length 5952;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 ACAATGGCAGACCTCCCAAGTGAATTAAGGTCACACAGATGTATTATCAACAACCTCGA 60	
Db	3102 ACATGGCGAGACTTCCAGGTGAAATTAAGGTCACACAGATGTATTATCAACAACCTCGA 316	
Qy	61 TGAAAACAGCCMAAAATCTTGAGATCCCTGGAAAGTTCCGATGATGCAGTCTCTGTATCA 120	
Db	3162 TGAAAACAGCCMAAAATCTTGAGATCCCTGGAAAGTTCCGATGATGCAGTCTCTGTATCA 322	
Qy	121 AAAGCGTTTGATTAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACAT 180	
Db	3222 AAGAGCTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACAT 328	
Qy	181 TAGGTCCCATTTGGAAAGCCAGTTCTGACCAGTGGAGCGTCTGCACCTTTCTCTGCAGGA 240	
Db	3282 TAGGTCCCATTTGGAAAGCCAGTTCTGACCAGTGGAGCGTCTGCACCTTTCTCTGCAGGA 334	
Qy	241 ACTTCTGGTGTGGCTTACAGCTGAAAGATGATGAANTAAAGCCGCCAGGCACCTATTTCGAGG 300	
Db	3342 ACTTCTGGTGTGGCTTACAGCTGAAAGATGATGAANTAAAGCCGCCAGGCACCTATTTCGAGG 340	
Qy	301 CGACTTTCCAGAGTTTCAGAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 360	
Db	3402 CGACTTTCCAGAGTTTCAGAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA 346	
Qy	361 AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGATGTACGGAATATTCTTCACAGAGCA 420	

Db	3462	AACATAAGAACCTGTAATCATAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA	3521
Qy	421	GCCTTTTGAAGAGACTAGAGAACTCTACAGGAGCCGAGAGAGCTGCCTCCTCAGGAGAG	480
Db	3522	GCCTTTTGAAGAGACTAGAGAACTCTACAGGAGCCGAGAGAGCTGCCTCCTCAGGAGAG	3581
Qy	481	AGCCAGAAATGCTACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATACTAGTGGGA	540
Db	3582	AGCCAGAAATGCTACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATACTAGTGGGA	3641
Qy	541	AAAAATCAACTGCTGACCTCCCTGACTGCGAGAGAAAATAGATGAGACCTTTGAAAGCT	600
Db	3642	AAAAATGAACTGCTGACCTCCCTGACTGCGAGAGAAAATAGATGAGACCTTTGAAAGCT	3701
Qy	601	CCAGGAATCTCAAGAGGCCACGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT	660
Db	3702	CCAGGAATCTCAAGAGGCCACGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT	3761
Qy	661	CAAGGGATCCTGGCAGCCGCTGGCGCATCTCCTCATTTGACTCTCTCCAAAGATCACTCGA	720
Db	3762	CAAGGGATCCTGGCAGCCGCTGGCGCATCTCCTCATTTGACTCTCTCCAAAGATCACTCGA	3821
Qy	721	GAAAGTCAAGGCATCTCGAGGAGAAAATTTGCCCTCTGAAAGAGAACTGAGCCAGCTCAA	780
Db	3822	GAAAGTCAAGGCATCTCGAGGAGAAAATTTGCCCTCTGAAAGAGAACTGAGCCAGCTCAA	3881
Qy	781	TGACCTTGTCTGCGCAGCTTACCACTTTGGGCATTTGAGCTCTCACCGTATAA	840
Db	3882	TGACCTTGTCTGCGCAGCTTACCACTTTGGGCATTTGAGCTCTCACCGTATAA	3941
Qy	841	TCCTGGAAGACCTGAAACACAGATGGAGCTTTCTGCAGGTGGCCCTCGAGGACCGAGTCA	900
Db	3942	TCCTGGAAGACCTGAAACACAGATGGAGCTTTCTGCAGGTGGCCCTCGAGGACCGAGTCA	4001
Qy	901	GCAGCTGCGATGAAGCCGACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC	960
Db	4002	GCAGCTGCGATGAAGCCGACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC	4061
Qy	961	TGTCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACA	1020
Db	4062	TGTCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACA	4121
Qy	1021	CGAGACTCAAAACAACTTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGC	1080
Db	4122	CGAGACTCAAAACAACTTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGC	4181
Qy	1081	TGACCTGGAATATGTACAGATTTCTAGCTTATAGGACTGCCATGAACCTCCGAAAGCTGCA	1140
Db	4182	TGACCTGGAATATGTACAGATTTCTAGCTTATAGGACTGCCATGAACCTCCGAAAGCTGCA	4241
Qy	1141	GAAAGCCCTTTGCTTGGATCTCTTGGAGCTGCTGAGCTGCTGATGCTTGGACCCAGCA	1200
Db	4242	GAAAGCCCTTTGCTTGGATCTCTTGGAGCTGCTGAGCTGCTGATGCTTGGACCCAGCA	4301
Qy	1201	CAACCTCAAGCAAAATGACAGCCCATGGATATCTGAGATTTAATTTGTTGACCAAC	1260
Db	4302	CAACCTCAAGCAAAATGACAGCCCATGGATATCTGAGATTTAATTTGTTGACCAAC	4361
Qy	1261	TATTTATGACCCCTGGAGAGAGCAACAATTTGGTCAACGTCCTCTCTCGCTGGA	1320
Db	4362	TATTTATGACCCCTGGAGAGAGCAACAATTTGGTCAACGTCCTCTCTCGCTGGA	4421
Qy	1321	TATGTCTGAACTGGCTGCTGAATGTTTATATACGGGACGAAACAGGGAGATCCCGTGT	1380
Db	4422	TATGTCTGAACTGGCTGCTGAATGTTTATATACGGGACGAAACAGGGAGATCCCGTGT	4481
Qy	1381	CCTGTCTTTTAAACTGGCATCATTTCCCTGTATAGACATTTTGGAGACAGTACAG	1440
Db	4482	CCTGTCTTTTAAACTGGCATCATTTCCCTGTATAGACATTTTGGAGACAGTACAG	4541
Qy	1441	ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCCAGGCTTGGGCCT	1500
Db	4542	ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCCAGGCTTGGGCCT	4601

Qy 1501 C 1501

Db 4602 C 4602

Search completed: September 19, 2004, 04:48:54
Job time : 413.027 secs

Db 3342 ACTTCTGGTGGCTACAGCTGAAGAATGATGAATTAAGCCGCGAGCACCCTATTGGAGG 3401
Qy 301 CGACTTTCCAGCAGTTCAGAAAGCAGAAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 360
Db 3402 CGACTTTCCAGCAGTTCAGAAAGCAGAAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 3461
Qy 361 AACTAAAGAACTGTATCATGATCTCTTGAGACTGTGACGAAATTTCTGACAGAGCA 420
Db 3462 AACTAAAGAACTGTATCATGATCTCTTGAGACTGTGACGAAATTTCTGACAGAGCA 3521
Qy 421 GCCTTTGGAGGACTCAGAGAACTCTACAGGAGCCAGAGAGCTGCCCTCAGGAGAG 480
Db 3522 GCCTTTGGAGGACTCAGAGAACTCTACAGGAGCCAGAGAGCTGCCCTCAGGAGAG 3581
Qy 481 AGCCAGAAATGTCTACCTCGGCTTTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 540
Db 3582 AGCCAGAAATGTCTACCTCGGCTTTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 3641
Qy 541 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAATAGATGAGACCTTGAAGACT 500
Db 3642 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAATAGATGAGACCTTGAAGACT 3701
Qy 601 CCAGGAATTTCAAGAGGCCACGGATGAGCTGAGCTCAAGCTCGGCCAAGCTGAGGTGAT 660
Db 3702 CCAGGAATTTCAAGAGGCCACGGATGAGCTGAGCTCAAGCTCGGCCAAGCTGAGGTGAT 3761
Qy 661 CAAGGGATCTGGCAGCCCGTGGCGATCTCTCATGATCTCTCTCAAGATCACTCGA 720
Db 3762 CAAGGGATCTGGCAGCCCGTGGCGATCTCTCATGATCTCTCTCAAGATCACTCGA 3821
Qy 721 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACTGAGCCACGCTCAA 780
Db 3822 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACTGAGCCACGCTCAA 3881
Qy 781 TGAACCTTGTGGCAGCTTACCACTTTGGGCACTGAGCTCTACCGTATAAAGCTCAGCAC 840
Db 3882 TGAACCTTGTGGCAGCTTACCACTTTGGGCACTGAGCTCTACCGTATAAAGCTCAGCAC 3941
Qy 841 TCTGGAAGACTCTGAACACAGATGGAAGCTTTCTGAGGTGGCGCTCGAGGACCGAGTCTAG 900
Db 3942 TCTGGAAGACTCTGAACACAGATGGAAGCTTTCTGAGGTGGCGCTCGAGGACCGAGTCTAG 4001
Qy 901 GCAGCTGATGAAGCCCAAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 960
Db 4002 GCAGCTGATGAAGCCCAAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 4061
Qy 961 TGTCCAGGCTCCCTGGAGAGAGCCATCTGGCCAAAGTGGCCCTACTATATCAACCA 1020
Db 4062 TGTCCAGGCTCCCTGGAGAGAGCCATCTGGCCAAAGTGGCCCTACTATATCAACCA 4121
Qy 1021 CGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGC 1080
Db 4122 CGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGC 4181
Qy 1081 TGACCTTGAATATGTGATGATCTCAGCTTATAGGACTGCCATGAACTCCGAGAGTGA 1140
Db 4182 TGACCTTGAATATGTGATGATCTCAGCTTATAGGACTGCCATGAACTCCGAGAGTGA 4241
Qy 1141 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGATGCTTGGAGCCAGCA 1200
Db 4242 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGATGCTTGGAGCCAGCA 4301
Qy 1201 CAACTTCAAGCAAAATGACAGCCCATGATATCTGAGAGTATTAATTTGTTGACCA 1260
Db 4302 CAACTTCAAGCAAAATGACAGCCCATGATATCTGAGAGTATTAATTTGTTGACCA 4361
Qy 1261 TATTTATGACCGCTGGAGCAAGAGCAAACTTTGTTGACCGTCTCTCTGCGTGA 1320
Db 4362 TATTTATGACCGCTGGAGCAAGAGCAAACTTTGTTGACCGTCTCTCTGCGTGA 4421
Qy 1321 TATGCTCTGAACCTGCTCTGAAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 1380
Db 4422 TATGCTCTGAACCTGCTCTGAAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 4481

Qy 1381 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAGCATTTTGAAGACAAGTACAG 1440
Db 4482 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAGCATTTTGAAGACAAGTACAG 4541
Qy 1441 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTACACAGCGAGGCTGGGCT 1500
Db 4542 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTACACAGCGAGGCTGGGCT 4601
Qy 1501 C 1501
Db 4602 C 4602

RESULT 2

US-09-484-970B-60

; Sequence 60, Application US/09484970B

; Patent No. 6426186

; GENERAL INFORMATION:

; APPLICANT: Jones, Karen A.

; APPLICANT: Voikmuth, Wayne

; APPLICANT: Walker, Michael G.

; TITLE OF INVENTION: BONE REMODELING GENES

; FILE REFERENCE: PB-0014 US

; CURRENT APPLICATION NUMBER: US/09/484,970B

; CURRENT FILING DATE: 2000-01-18

; NUMBER OF SEQ ID NOS: 172

; SOFTWARE: PERL Program

; SEQ ID NO 60

; LENGTH: 13977

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1

; NAME/KEY: unsure

; LOCATION: 11721-11761, 12294, 13969

; OTHER INFORMATION: a, t, c, g, or other

US-09-484-970B-60

Query Match 99.8%; Score 1497.4; DB 4; Length 13977;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1498; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACATGTCAGAGACTCCAAAGTGAATTTGAAGTCTCAGAGATGTTTATCAGAACCTGGA 60
Db 8416 ACATGTCAGAGACTCCAAAGTGAATTTGAAGTCTCAGAGATGTTTATCAGAACCTGGA 8475
Qy 61 TGAAGACAGCCAAAATCCTGAGATCCCTGGAAAGTTCGATGATGCAAGTCTCTGTACA 120
Db 8476 TGAAGACAGCCAAAATCCTGAGATCCCTGGAAAGTTCGATGATGCAAGTCTCTGTACA 8535
Qy 121 AAGAGTTTGGATTAACATGAACCTTCAAGTGAAGTGAATTCGGAAAAAGTCTCTCAACAT 180
Db 8536 AAGAGTTTGGATTAACATGAACCTTCAAGTGAAGTGAATTCGGAAAAAGTCTCTCAACAT 8595
Qy 181 TAGTCCCATTTTGGAAAGCCAGTTCTGACAGTGGAAAGTTCGACCTTCTCTGAGCA 240
Db 8596 TAGTCCCATTTTGGAAAGCCAGTTCTGACAGTGGAAAGTTCGACCTTCTCTGAGCA 8655
Qy 241 ACTTCTGGTGGCTCAGCTGAAAGATGATGAATTAAGCCGAGGACCTATTGGAGG 300
Db 8656 ACTTCTGGTGGCTCAGCTGAAAGATGATGAATTAAGCCGAGGACCTATTGGAGG 8715
Qy 301 CGACTTTCCAGCAGTTCAGAGCAGAACCATGTGTACATAGGCTTCAAGAGGGAATTGAA 360
Db 8716 CGACTTTCCAGCAGTTCAGAGCAGAACCATGTGTACATAGGCTTCAAGAGGGAATTGAA 8775
Qy 361 AACTAAAGAACTGTATCATGATCTCTTGAGACTGTGACGAAATTTCTGACAGAGCA 420
Db 8776 AACTAAAGAACTGTATCATGATCTCTTGAGACTGTGACGAAATTTCTGACAGAGCA 8835
Qy 421 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTCCCTCTCTGAGGAG 480

QY 361 AACTAAAGAACCTGTATCATATGAGTACTCTGTGAGACGTGTACGAATATTTCTGTACAGAGCA 420
Db 5918 AACTAAAGAACCTGTATCATATGAGTACTCTGTGAGACGTGTGTGAGATATTTCTGTACAGAGCA 5859
QY 421 GCCTTTTGAAGGACTAGAGAACTCTACACAGGAGCCGACAGAGCTGCCTCCTGAGGAGAG 480
Db 5858 GCCTTTTGAAGGACTAGAGAACTCTACACAGGAGCCGACAGAGCTGCCTCCTGAGGAGAG 5799
QY 481 AGCCAGAAATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATCACTGAGTGGGA 540
Db 5798 AGCTCAGAATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATCACTGAGTGGGA 5739
QY 541 AATATTGAACCTGCACTCCGCTGACTGCTGGCAGAGAAAATAGATGACACCTTGAAGACT 600
Db 5738 CAATTTGAACCTGCTGCTGACTGCTGGCAGAGAAAATAGATGAGACTTCTTGAAGACT 5679
QY 601 CCAGGAATCTCAAGAGCCACGATGAGCTGAGCTGAGCTGCGCAAGCTGAGGTGAT 660
Db 5678 CCAGGAATCTCAAGAGCTGCGGATGAGCTGAGCTGCGCAAGCTGAGGTGAT 5619
QY 661 CAAGGATCTCGGAGCCCTGCGGAGATCTCTCATATGACTCTCTCAAGATCACTGCA 720
Db 5618 CAAGGATCTCGGAGCCCTGCGGAGATCTCTCATATGACTCTCTCAAGATCACTGCA 5559
QY 721 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACTGAGCCAGCTCAA 780
Db 5558 AAGAGTCAAGGCACTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACTGAGCCAGCTCAA 5499
QY 781 TGACCTTGCTGCGCAGCTTACCACTTTGGCAGTTCAGCTCTCACGATTAACCTCAGCAC 840
Db 5498 TGACCTTGCTGCGCAGCTTACCACTTTGGCAGTTCAGCTCTCACGATTAACCTCAGCAC 5439
QY 841 TCTGGAAGACTCTGAACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGAGCGAGTCA 900
Db 5438 TTTGGAAGACTCTGAATACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGAGCGAGTCA 5379
QY 901 GCAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCATCTTCTTTCCACAGTC 960
Db 5378 ACAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCATCTTCTTTCCACAGTC 5319
QY 961 TGTCAGGCTCCCTGGAGAGAGCCATCTGCGCAAAAGAGTCCCTACTATATCAACCA 1020
Db 5318 AGTTCAGGCTCCCTGGAGAGAGCCATCTGCGCAAAAGAGTCCCTACTATATCAACCA 5259
QY 1021 CGAGACTCAAAACAACTTGTCTGGAGACCATCCCAAAATGACAGAGCTCTACAGCTCTTTAGC 1080
Db 5258 CGAGACTCAAAACAACTTGTCTGGAGACCATCCCAAAATGACAGAGCTCTACAGCTCTTTAGC 5199
QY 1081 TGACCTGAATATGTGAGATCTCAGCTTATAGGATGCGCATGAAACTCCGAGAGCTGCA 1140
Db 5198 TGACCTGAATATGTGAGATCTCAGCTTATAGGATGCGCATGAAACTCCGAGAGCTGCA 5139
QY 1141 GAAGGCCCTTTGCTGATCTCTGAGCTGTGAGCTGATGATGCTGAGAGAGCA 1200
Db 5138 GAAGGCCCTTTGCTGATCTCTGAGCTGTGAGCTGATGATGCTGAGAGAGCA 5079
QY 1201 CAACCTCAAGCAAAATGACAGCCCATGATATCTGCGAGATTAATTAATTTGTTGACCA 1260
Db 5078 CAACCTCAAGCAAAATGACAGCCCATGATATCTGCGAGATTAATTAATTTGTTGACCA 5019
QY 1261 TATTTATGACCGCTGGAGAGAGCAACAATTTGTTGAGCTCCCTCTCTGCGTGA 1320
Db 5018 AATTTATGATGCTGGAGAGAGCAACAATTTGTTGAGCTCCCTCTCTGCGTGA 4959
QY 1321 TATGCTCTCAACTGCTCTCAATGTTTATGATACCGGAGCAACAGGAGGATCCGCTGT 1380
Db 4958 TATGCTCTCAACTGCTCTCAATGTTTATGATACCGGAGCAACAGGAGGATCCGCTGT 4899
QY 1381 CTTGCTCTTTAAACTGGGATCATTTCTCTGTGTAAGGACATTTTGAAGAGCAAGTACAG 1440
Db 4898 CTTGCTCTTTAAACTGGGATCATTTCTCTGTGTAAGGACATTTTGAAGAGCAAGTACAG 4839

QY 1441 ATACTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACAGCGCAGGCTGGGCT 1500
Db 4838 ATACTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACAGCGCAGGCTGGGCT 4779

RESULT 4

US-09-427-048A-10/c
; Sequence 10, Application US/09427048A
; Patent No. 6203975
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; Wilson, James M.
; Fisher, Krishna J.
; Chen, Shu-Jen
; Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; Methods of Use Thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/427,048A
; FILING DATE: 21-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,022
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNVN.008PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19307 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-427-048A-10

Query Match 86.5%; Score 1298.4; DB 3; Length 19307;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
QY 1 ACAATGCAAGACCTCCAGGTGAAATTTGAAGTCAACAGATGTTTATCAACACCTGGA 60
Db 6278 ACATGCAAGATCTCCAGGAGAAATTTGAATCAACAGATATCTATCAACATTTGA 6219
QY 61 TGAACAGACCAAAATCTGAGATCCCTGGAAGGTTCCGATGATGAGTCCGTTACA 120
Db 6218 TGAATGGAACCAAAATCTGAGATCCCTGGAAGGTTCCGATGAGTCCGTTACA 6159
QY 121 AAGACGTTTGATAACATGAATTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACAT 180
Db 6158 AAGACGTTTGATAACATGAATTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACAT 6099
QY 181 TAGGTCCCAATTTGGAAGCAGGTCTGACAGTGAAGAGCTCTGACCTTTCTCTCAGGA 240
Db 6098 TAGGTCCCAATTTGGAAGCAGGTCTGACAGTGAAGAGCTTTGCAATCTTCTCTCAGGA 6039

QY 241 ACTTCTGTGGTGTGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTTATTGGAGG 300
Db 6038 ACTTCTGTGGTGTGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTTATTGGAGG 5979
QY 301 CGACTTTCAGCAGTTCAGAACGAGACGATACATAGGGCTTCAAGAGGGAATTGAA 360
Db 5978 TGATTTCCAGCAGTTCAGAACGAGATGATATACATAGGGCTTCAAGAGGGAATTGAA 5919
QY 361 AACTAAAGAACCTGTAATCATGAGTACTTTCAGAGCTGTACGAATATTTCTGACAGACA 420
Db 5918 AACTAAAGAACCTGTAATCATGAGTACTTTCAGAGCTGTGAGAAATTTCTGACAGACA 5859
QY 421 GCCTTTGGAAGACTAGAAACTCTTACCAGAGGCCAGAGAGCTGCCTCTCTGAGGAGAG 480
Db 5858 GCCTTTGGAAGACTAGAAACTCTTACCAGAGGCCAGAGAGCTGCCTCTCTGAGGAAAG 5799
QY 481 AGCCAGAAATGCACTCGGCTTCTAGCAAGCAGGCTGAGGAGTCAATACCTGAGTGGGA 540
Db 5798 AGCTCAGAATGCACTCGGCTTCTAGCAAGCAGGCTGAGGAGTCAATACCTGAGTGGGA 5739
QY 541 AAAATTGAACTGCACTCGGCTTCTAGCAAGCAGGCTGAGGAGTCAATACCTGAGTGGGA 600
Db 5738 CAAATTGAACTGCACTCGGCTTCTAGCAAGCAGGCTGAGGAGTCAATACCTGAGTGGGA 5679
QY 601 CCAGGAACCTTCAAGAGCCCAAGATGAGTGGAGCTCAAGCTGCGCAAGCTGAGGTGAT 660
Db 5678 CCAGGAACCTTCAAGAGCTGCGCAAGATGAGTGGAGCTCAAGCTGCGCAAGCTGAGGTGAT 5619
QY 661 CAAGGATCTGCGAGCCGCTGAGGCTGAGTGGAGCTCAAGCTGCGCAAGCTGAGGTGAT 720
Db 5618 CAAGGATCTGCGAGCCGCTGAGGCTGAGTGGAGCTCAAGCTGCGCAAGCTGAGGTGAT 5559
QY 721 GAAAGTCAAGGCACTTCGAGGAGAAATGCGCTCTGAAAGAGAAAGCTGAGGACCTCAA 780
Db 5558 AAGGTCAGGCACTTCGAGGAGAAATGCGCTCTGAAAGAGAAAGCTGAGGACCTCAA 5499
QY 781 TGACCTTGTGCGCAGCTTACCACTTTGGCATTCAGCTCTACCGTATAACTCAGCAGAC 840
Db 5498 TGACCTTGTGCGCAGCTTACCACTTTGGCATTCAGCTCTACCGTATAACTCAGCAGAC 5439
QY 841 TCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAAGTGGCGCTGAGGACCGAGTCA 900
Db 5438 TTTGGAAGATCTGAATACCAAGATGAGGCTTCTACAGGTGGCTGTGAGGACCGTGTCA 5379
QY 901 GCAGCTGCATGAAGCCACAGGACCTTTGCTCAGCATCTCAGCATCTTCTTTCACGTC 960
Db 5378 ACAGCTGCATGAAGCCACAGGACCTTTGCTCAGCATCTCAGCATCTTCTTTCACGTC 5319
QY 961 TGTCAGGCTCCCTGGGAGAGAGCACTCTCGCAAAACAAAGTCCCTACTATATCAACCA 1020
Db 5318 AGTTCCAGGCTCCCTGGGAGAGAGCACTCTCGCAAAACAAAGTCCCTACTATATCAACCA 5259
QY 1021 CGAGACTCAAAACACTTGTGCTGGGACCAATCCCAAAATGACAGACTCTACAGTCTTTAGC 1080
Db 5258 CGAGACCCAAACACTTGTGCTGGGACCAATCCCAAAATGACAGACTCTACAGTCTTTAGC 5199
QY 1081 TGACCTGAATATGTCAGATCTCAGTGTATAGGACTGCCATGAACTCCGAGACTGGA 1140
Db 5198 TGACCTGAATATGTCAGTGTATAGGACTGCCATGAACTCCGAGACTGGA 5139
QY 1141 GAGGCTCTTGTGCTGATCTTGTAGCTGTGATGCTGATGCTGATGCTGATGCTGATGCT 1200
Db 5138 GAGGCTCTTGTGCTGATCTTGTAGCTGTGATGCTGATGCTGATGCTGATGCTGATGCT 5079
QY 1201 CAACCTCAAGCAAAATGACAGGCCATGATATCTCTGAGATTAATTTGTTGACAC 1260
Db 5078 CAACCTCAAGCAAAATGACAGGCCATGATATCTCTGAGATTAATTTGTTGACAC 5019
QY 1261 TATTTATGACCGCTGAGGAGAGACACAAATTTGGTCAAGCTCCCTCTCTGCTGTGA 1320
Db 5018 AATTTATGATCGCTGAGGAGAGAGACACAAATTTGGTCAAGCTCCCTCTCTGCTGTGA 4959

QY 1321 TATGTGTCTGAATGCTGCTGAATGTTTATGATAGCGGACGAAACAGGAGGATCCGTGT 1380
Db 4958 TATGTGTCTGAATGCTGCTGAATGTTTATGATAGCGGACGAAACAGGAGGATCCGTGT 4899
QY 1381 CTTGTCTTTTAAACTGGGATCAATTTCCCTGTCTAAAGACATTTTGAAGACAAGTACAG 1440
Db 4898 CTTGTCTTTTAAACTGGGATCAATTTCTGTGTAAAGACAATTTGAAGACAAGTACAG 4839
QY 1441 ATACCTTTTCAAGCAGAGTGGCAAGTTTCAACAGGATTTTGTACAGCGCAGGTGGGCT 1500
Db 4838 ATACCTTTTCAAGCAGAGTGGCAAGTTTCAACAGGATTTTGTACAGCGCAGGTGGGCT 4779
RESULT 5
US-09-091-501B-7
; Sequence 7, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathan M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091.501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 6045
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(6037)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; NAME/KEY: misc feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
; US-09-091-501B-7
Query Match 40.3%; Score 604.8; DB 4; Length 6045;
Best Local Similarity 63.1%; Pred. No. 1.1e-187;
Matches 952; Conservative 0; Mismatches 547; Indels 9; Gaps 1;
QY 2 CAAATGGCAGACCTTCCCAAGGTGAAATTCAGCTCAGAGATGTTTATCAACACCTGGAT 61
Db 3206 CAGATGCAAGGACATCCAGGAGAAATTCATGCCCAATGACATATTTAAAGCATTCAC 3265
QY 62 GAAACACGCAAAATTCCTGAGATCCCTGGAAGTTCCGATGATGATGATGATGATGAT 121
Db 3266 GGAACACGCAAGATGTTAAAGCTTTGGAATTTCTGAAGAGGCTTACTATGCTTCAA 3325
QY 122 AGACCTTTGATTAACATGAATTCAGTGGAGTGAATTCGGAAGAAAGTCTCTCAACATT 181
Db 3326 CATGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3385
QY 182 AGGTCCTTGGAGGAGGAGTTCGACAGTGGAGGCTCTGCACTTTCTCTGAGGAA 241
Db 3386 AGGTCCTTGGAGGAGGAGTTCGAGAGTGGAGGAGTTCGAGGAGTTCGAGGAGTTCGAG 3445
QY 242 CTTCTGGTGTGCTACAGTGAAGATGATGAATTAAGCCGCGAGGACCTTATTGGAGGC 301
Db 3446 CTGATCAATGCTGAATGATGAAGATGATGAAGCTTAAAGAAACAAATCCCTATTGGAGG 3505

QY 302 GACITTCACGAGTTCAGAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAAA 361
Db 3506 GATGTTCCAGCTTACAGCTCCAGTATGACATTTGAAGGCCCTGAGACGGAGTTAAAG 3565
QY 362 ACTAAGAACCTGTAATCATGAGTACTCTTGAAGACTGTACGAATATTTCTGACAGAGCAG 421
Db 3566 GAGAAAGAAATATTCTGCTCTGAAATGCTGTGACAGGCCGAGTTTCTTGCTGTATCAG 3625
QY 422 CTTT-----GGAGACTAGAGAACTCTACAGAGGCCACAGAGCTGCTCTCT 472
Db 3626 CCAATTGAGGCCCTGAGAGCCCAAGAGAACTCTCAATCAAAAAACAGAAATTAACCTCT 3685
QY 473 GAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTCAGAGGTCATATCT 532
Db 3686 GAGGAGAGAGCCCAAGAGTTGCCAAGCCATCGGCAACAGTCTTCTGAAGTCAAGAA 3745
QY 533 GAGTGGAAAAATGAACTGCACTCGCTGACTCGGCTGACTGCGCAGAGAAAAATAGATGAGACCTT 592
Db 3746 AAATGGGAAAGTCTAAATGCTGTAACTAGCAATTTGCAAAAGCAAGTGACAAAGGCATTG 3805
QY 593 GAAAGACTCCAGGAATCTCAAGAGGCCACGATGCTGACCTCAAGCTGGCCAGCT 552
Db 3806 GAGAACTCAGACCTTGCAGGAGCTATGATGACTGAGCCTGACATGAGAGGAGCA 3865
QY 653 GAGTGATCAAGGATCTTGGCAGCCCGTGGGCGATCTCTATTGACTCTCTCAAAGAT 712
Db 3866 GAGTCCGTGGGAATGGCTGGAAGCCCGTGGGAGACTTACTATTGACTCGCTGAGGAT 3925
QY 713 CAGCTCGAGAACTCAAGCACTTCAGAGAGAAATTTGGCCTCTGAAAGAGAACTGAGC 772
Db 3926 CAGTTGAAAAAATCATGGCAATTTAGAGAGAAATTTGCACCAATCACTTTAAAGTTAA 3985
QY 773 CAGCTCAATGACCTTGTGCTGCCAGCTTACCACTTTTGGGCACTCAGCTCTCAACGCTATAAC 832
Db 3986 ACGTGAATGATTTATCCAGTCACTGTCTCCACTTGACTGCATCCCTCTTAAGATG 4045
QY 833 CTGACACTCTGAAGACCTGAACACAGATGAAGCTTCTGAGGTGGCCCTGAGAGAC 892
Db 4046 TCTCGCCAGCTAGATGACCTTAAATATGTCGATGGAACCTTTACAGGTTTCTGTGATGAT 4105
QY 893 CGAGTCAGGAGCTGATGAAGCCACAGGACTTTGTCCAGCATCTCAGACATTTCTT 952
Db 4106 CGCTTAAACAGTTTCAAGAGCCACAGATTTTGGACCATCTCTCAGCATTTTCTC 4165
QY 953 TCACGCTGTCCAGGTCCTGGGAGAGAGCACTCTGCCAAACAAAGTGCCCTACTAT 1012
Db 4166 TCTACGTCAGTCCAGCTGCGTGGCAAGATCCATTTACATAATAAAGTGCCCTATTAC 4225
QY 1013 ATCAACCCAGAGACTCAACCACTTGTCTGGGAGCCATCCCAAAATCAGAGCTCTACCAG 1072
Db 4226 ATCAACCATCAACACAGACCACTTGTGGGACCATCTTAAATGACCGAATCTTTCAA 4285
QY 1073 TCTTTAGCTGACCTGAATATGTCAAGTTCTCAGGTTATAGGACTGCCATGAACTCCGA 1132
Db 4286 TCCCTTGTGACTGAAATATGTAAGTTTCTGCTACCGTACAGCAATCAAAATCCGA 4345
QY 1133 AGACTCAGAGGCCCTTTGCTTGTATCTTTGAGCTGTGAGTGTGATGCTGCTT 1192
Db 4346 AGACTCAAAAAGCACTATGTTGGATCTTTAGAGTTTGAATACCAAAATGAAATTTTC 4405
QY 1193 GACGAGCAACCTCAAGCAAAATACAGGCCATCGGATATCTCAGATTTAATTTGT 1252
Db 4406 AAACACCAAGTTGAAACCAAAATGACAGCTCTCTGAGTTCCAGATGTCTCACTGT 4465
QY 1253 TTGACCACTATTATGACCGCTTGAGCAAGAGCAACAATTTGGTCAAGTCCTCTC 1312
Db 4466 CTGACAACTATTATGATGACTTGAGCAAAATGCAATAAGGACCTGCTCAAGCTTCACTC 4525
QY 1313 TGGTGGATATGTCTGAACTGGCTGTGTAATTTTATGATACGGGAGCAACAGGAGG 1372
Db 4526 TGTGTTGATATGTCTCAATTTGGTTGCTCAATGCTATGACCGGGTGCAACTGGAAAA 4585
QY 1373 ATCCGTGCTCTGCTTTTAAAACTGGCATCATTTTCCCTGTGTGTAAGACACATTTGGAAGAC 1432

RESULT 6

US-09-091-501B-9
; Sequence 9, Application US/09091501B

; Patent No. 6518413

; GENERAL INFORMATION:

; APPLICANT: Tinsley, Jonathon M

; APPLICANT: Davies, Kay E

; TITLE OF INVENTION: Utrophin gene expression

; FILE REFERENCE: 620-42

; CURRENT APPLICATION NUMBER: US/09/091,501B

; CURRENT FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: PCT/GB96/03156

; PRIOR FILING DATE: 1996-12-19

; PRIOR APPLICATION NUMBER: GB 9525962.8

; PRIOR FILING DATE: 1995-12-19

; PRIOR APPLICATION NUMBER: GB 9615797.9

; PRIOR FILING DATE: 1996-07-26

; PRIOR APPLICATION NUMBER: GB 9622174.2

; PRIOR FILING DATE: 1996-10-24

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 9

; LENGTH: 10320

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (11)..(10312)

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Full length

; OTHER INFORMATION: utrophin construct

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (724)..(758)

; OTHER INFORMATION: Precise residue is left open

US-09-091-501B-9

Query Match 40.3%; Score 604.8; DB 4; Length 10320;

Best Local Similarity 63.1%; Pred. No. 1.5e-187;

Matches 952; Conservative 0; Mismatches 547; Indels 9; Gaps 1;

QY 2 CAATGCGACACCTCCAAAGTGAATTTGAAGTTCACACAGATGTTTATCACACCTGGAT 61
Db 7481 CAGATGAGACATCCAGGAGGAGAAATTTGATGCCCAATGACATATTTAAAGCATTTGAC 7540
QY 62 GAAACACAGCCAAAAATTCCTGAGATCCCTGGAGGTTCCGATGATGAGTCTCTGTACAA 121
Db 7541 GMAACAGGCGAAGATGTTAAAGCTTTGGGAAATTTGAAAGAGGCTACTATGCTTCAA 7600
QY 122 AGAGCTTGGATGAATGAATTTCAAGTGAAGTGAAGTTCGGAAGTCTCTCAACATT 181
Db 7601 CATCGACTGGATGATGATGAACCAAGATGGAATGACTTAAAGCAAAATCTCTAGCATC 7660
QY 182 AGGTCCCATTTGGAAGCCAGTCTGACCACTGGAAGCGTCTGACCTTTCTCTGAGGAA 241
Db 7661 AGGGCCATTTGGAGGCCAGCGCTGAGAAGTGAAGAGTCTCTGATGCTCTTAGAGAA 7720
QY 242 CTTCTGGTGGCTACAGCTGAAAGTGAATTAAGCCCGCAGGACCTATTATGGAGCC 301
Db 7721 CTGATCAAAATGGCTGAATGATGAAGATGAAGAGCTTAAGAAAACAAATGCCTATTGGAGGA 7780

Db 1022 GGGCATGGAGAACTAAGCACTACTCTGAGCCAGCTGAGGGAGTCCGAGCCACTTTGGGA 1081
 QY 576 GCGCGTGGGAGTCTCTCACTTGAATCTCTCAAGATCACTCGAGAAAGTCAAGCACT 735
 Db 1082 GCGCATTTGGGATCTCTTCACTTGAATCACTCGAGAGCAATCCAGGCTATTAAAGTGT 1141
 QY 736 TCGAGGAGAAATTCGGCTCTGAAAGAGAACTGAGCCAGTCAATGACCTTGTCTGCCA 795
 Db 1142 CAAAGAGAAATTCCTCCCATGAAGATGAGTAAAGTGGTGAATGATCTGCGCCACCA 1201
 QY 796 GCTTACCACTTTGGGCAATTCAGCTCTCACCTTAACCTCAAGCACTCTGGAAGACCTGAA 855
 Db 1202 ACTTGCCATTTCTGATGTGCACTTGTCAATGGAGAAATTCAGGCGCTCGAAGACATCA 1261
 QY 856 CACAGATGAAGCTTCTGAGGTGCGCGTCCAGGAGCCGAGTCCAGGAGCTGCATGAGC 915
 Db 1262 CTGCGATGGAACAATCACTAGCGCTCAGTTGATGAGGCTTAAGCAGCTCCAGGATCC 1321
 QY 916 CCACAGGACTTTGGTCCAGCACTCTCAGCACTTTCTTCCAGCTCTGTCCAGGCTCCCTG 975
 Db 1322 CCACCGGACTTTGGGCTGGGTCAAGCACTTTCTCTCTCTGTCCAGGTTCCCTG 1381
 QY 976 GGAGAGAGCACTCGCCCAACAAAGTGCCTACTATATCAACCCAGGAGCTCAACAAAC 1035
 Db 1382 GGAAGAGCAATTTCAACCAATTAAGTTCCTACTACATCAACCCAGGCTCAGACCAAC 1441
 QY 1036 TTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAAGT 1095
 Db 1442 ATGCTGGACCATCCCAAGATGACAGATTAACCAACCTAGCTGATCTGAAACAAT 1501
 QY 1096 CAGATCTCAGCTTATAGGACTGCCATGAACTCGAAGACTGAGAGGCGCTTTGCTT 1155
 Db 1502 TAAGTTCTCAGCTTATCGCACTGCCATGAACTCCGAGAGTCCGAGAAAGCGCTGCGCT 1561
 QY 1156 GGATCTCTTGAGCTGTGAGCTGATGATGCTTGGACCCAGCACCACTCAAGCAAA 1215
 Db 1562 GGACTGTGTAATTTAACCAAGCCCTGGAATCTTCAATGAGCATGATCTCGAGCCAG 1621
 QY 1216 TGACAGCCCATGATATCTCGAGATTAATTTGTTGACCACTATTATGACCGCT 1275
 Db 1622 TGAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1681
 QY 1276 GGAGCAGAGCACAAATTTGTCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1335
 Db 1682 GGAGGAGAAAGGAGCTCTGTCAGCTGCTGTCAGCTGCTGTCAGCTGCTGTCAGCTGCT 1741
 QY 1336 GCTGCTGAATGTTTATGATACGGGAGCAACAGGAGGATCCGCTGCTCTCTCTCTCTCT 1395
 Db 1742 GCTCCTCAATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1801
 QY 1396 TGGCATCATTTCCCTGTTGAAGCAGATTTGGAAGCAGATGATGATGATGATGATGATGAT 1455
 Db 1802 TGGCATGATGCTGTTGTGGCAGGAGTGAAGGAGAAACTTCACTACCTCTTCAGCCA 1861
 QY 1456 AGTGCAAGTCTCAACAGGATTTTGTGACCCAGGAGGCTGGGCGCTC 1501
 Db 1862 AGTGCCCACTCAGGAGCAGTGTGACCCAGGAGGCTGGGCGCTC 1907

RESULT 8
 ; US-09-687-875A-13
 ; Sequence 13, Application US/09687875A
 ; Patent No. 6544786
 ; GENERAL INFORMATION:
 ; APPLICANT: Xiao, Xiao
 ; APPLICANT: Liu, Paul
 ; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPICED PE
 ; FILE REFERENCE: 00792
 ; CURRENT APPLICATION NUMBER: US/09/687,875A
 ; CURRENT FILING DATE: 2000-10-13
 ; PRIOR APPLICATION NUMBER: 60/158,868
 ; PRIOR FILING DATE: 1999-10-15

; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 13
 ; LENGTH: 238
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: pXX-C2 5' junction
 ; US-09-687-875A-13
 Query Match 4.2%; Score 63.6; DB 4; Length 238;
 Best Local Similarity 94.3%; Pred. No. 9.4e-11;
 Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 97 TTCGATGATGTCAGTCTCTGTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGA 156
 Db 169 TTCGAGCAGCAGTACTGTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGA 228
 QY 157 ACTTCGGAAA 166
 Db 229 ACTTCGGAAA 238

RESULT 9
 ; US-08-232-463-14/c
 ; Sequence 14, Application US/08232463
 ; Patent No. 5670367
 ; GENERAL INFORMATION:
 ; APPLICANT: DORNER, F.
 ; APPLICANT: SCHEIFLINGER, F.
 ; APPLICANT: FALKNER, F. G.
 ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 ; NUMBER OF SEQUENCES: 52
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 1800 Diagonal Road, Suite 500
 ; CITY: Alexandria
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22313-0299
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/232,463
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/935,313
 ; FILING DATE:
 ; APPLICATION NUMBER: EP 91 114 300.6
 ; FILING DATE: 26-AUG-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BENT, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703)836-9300
 ; TELEFAX: (703)683-4109
 ; TELEX: 899149
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 7218 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; CLONE: PTZ9pt-Fls
 ; US-08-232-463-14

Query Match 3.6%; Score 53.4; DB 1; Length 7218;

APPLICANT: Hinman, Michael B.
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSES: Birch, Stewart, Kolasch & Birch
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22046
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/317,844B
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-105P
TELEPHONE: (703) 241-1300
TELEFAX: (703) 241-2848
TELEX: 248345
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1995 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
IMMEDIATE SOURCE:
CLONE: pS8
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1785
US-08-317-844B-3

Query Match 2.6%; Score 38.6; DB 2; Length 1995;
Best Local Similarity 44.0%; Pred. No. 0.067;
Matches 164; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY 311 GCAGTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAACTAAAGAA 370
DB 731 GCAGTCAGCTGAGCAGCGCCGAGGAGGACCTGGACACAAAGACCCGAGGATATGGAC 790

QY 371 CCTGTATATCATAGTACTCTTGAGACTGTAGAAATATTTCTGACAGAGCAGCCTTTGGAA 430
DB 791 CAGGACAAACAGGACCATTCTGGACCCCGGTAGTCCCGCTGACAGCAGCGCCGCGAGCAG 850

QY 431 GGACTAGAGAACTTACTACGAGGAGCCAGAGAGCTGCTCCTGAGGAGAGAGCCAGAAAT 490
DB 851 GACCTGGAGATATGGCCCTTGGACACAAAGACCCGAGGATATGACCCAGGACACAG 910

QY 491 GTCACCTGGCTTCTAAGAAAGCAGGCTGAGAGGTCATATCTGAGTGGGAAAAATTTGAAC 550
DB 911 GACCATCTGGAGCAGGCGAGTGCAGCAGCAGCGCCGAGGAGCCTTGGACACAAAGGAT 970

QY 551 CTGCACTCCCGCTGACTGGCAGAGAAATAGATGAGACCCCTTGAAGACTCTCCAGGAATTT 610
DB 971 TAGGAGGTTATGACCCAGGACACAAAGGTCAGAGGATATGGACCCAGGACAAACAGGTC 1030

QY 611 CAAGAGGCCCGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCC 670
DB 1031 CAGAGGATATGGACCCAGGATGTCATCTGCAGCAGCAGCGCCGAGGAGCCTGGACAAC 1090

QY 671 TGGAGCCCGTGG 683

Db 1091 AAGGACCAGGAGG 1103

RESULT 15
US-09-220-132-24
; Sequence 24, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANC.
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 7672
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-220-132-24

Query Match 2.6%; Score 38.4; DB 4; Length 7672;
Best Local Similarity 47.2%; Pred. No. 0.19;
Matches 117; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 675 AGCCCGTGGCGATCTCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCAC 734
DB 6431 ACCCGTGGCTGCAACTCTTGGAGAAATCAAAGCTTTGGCGAGGCCACACGCCT 6490

QY 735 TTCAGGAGAAATTGGCGCTCTGAAAGAGAACGTGACCCAGTCAATGACCTTCTCGCC 794
DB 6491 TCCGCTCTCTCCCTCAGCTCTGCCCGAGGTGACTTCAACCGAGCTGGCCGAGCTGACCGCC 6550

QY 795 AGCTTACCACCTTTGGGCACTCAGCTCTCACGTATAACCTCAGCACTCTGGAAGACCTGA 854
DB 6551 AGATCAAGAGCTTCCGGGTAGCTTCCAAACCCCTACACCTGGTTTACCATGGAGGCCCTGG 6610

QY 855 ACACCAAGATGGAAGCTTTCTGAGGTGGCCGCTGAGAGCCGAGTACGAGCAGTGCATGAAG 914
DB 6611 AGGAGACCTGAGGAACCTACAGAAATCATCAAGAGAGGAGGAGCTGGAGCTGCAGAAGG 6670

QY 915 CCCACAGG 922
DB 6671 AACAGCGG 6678

Search completed: September 19, 2004, 11:06:55
Job time : 81.4052 secs

Page Blank (uspiu)

181 TAGTCCCATTTGGAAGCCAGTCTTGACCAAGTGGAGCGTCTGCACCTTTCTCTGCAGGA 240
Db |||||
331 TAGTCCCATTTGGAAGCCAGTCTTGACCAAGTGGAGCGTCTGCACCTTTCTCTGCAGGA 390
Qy |||||
241 ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGCACCTATTGGAGG 300
Db |||||
391 ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGCACCTATTGGAGG 450
Qy |||||
301 CGACTTTCCAGCAGTTCAGAAAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 360
Db |||||
451 CGACTTTCCAGCAGTTCAGAAAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 510
Qy |||||
361 AACTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 420
Db |||||
511 AACTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA 570
Qy |||||
421 GCCTTTGGAAGAGCTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCTCTCTGAGAGAG 480
Db |||||
571 GCCTTTGGAAGAGCTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCTCTCTGAGAGAG 630
Qy |||||
481 AGCCCAAGATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 540
Db |||||
631 AGCCCAAGATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 690
Qy |||||
541 AAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 600
Db |||||
691 AAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 750
Qy |||||
601 CCAGGAACTTCAGAGGCCAGGAGCTGAGCTGAGCTCAAGCTCGGCCAAGCTGAGGTGAT 660
Db |||||
751 CCAGGAACTTCAGAGGCCAGGAGCTGAGCTGAGCTCAAGCTCGGCCAAGCTGAGGTGAT 810
Qy |||||
661 CAAGGGATCTCGCAGCCGCTGGGCGATCTCTCTCATTTGACTCTCTCCAAAGATCACTCGA 720
Db |||||
811 CAAGGGATCTCGCAGCCGCTGGGCGATCTCTCTCATTTGACTCTCTCCAAAGATCACTCGA 870
Qy |||||
721 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCGCTCTGAAAGAGACGCTGAGCCAGTCAA 780
Db |||||
871 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCGCTCTGAAAGAGAGACGCTGAGCCAGTCAA 930
Qy |||||
781 TGACCTTGCTCGCAGCTTACCACTTTGGGCACTTCAGCTCTCACCGTATAACCTCAGAC 840
Db |||||
931 TGACCTTGCTCGCAGCTTACCACTTTGGGCACTTCAGCTCTCACCGTATAACCTCAGAC 990
Qy |||||
841 TCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGAGAGTGGCGCTCGAGGACCGAGTCA 900
Db |||||
991 TCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGAGAGTGGCGCTCGAGGACCGAGTCA 1050
Qy |||||
901 GCAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 960
Db |||||
1051 GCAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 1110
Qy |||||
961 TGTCCAGGCTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCCTACTATATCAAGCA 1020
Db |||||
1111 TGTCCAGGCTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCCTACTATATCAAGCA 1170
Qy |||||
1021 CGAGCTCAAAACACTTGTGGGACCATCCCAAATGACAGAGCTCTACCACTCTTTAGC 1080
Db |||||
1171 CGAGCTCAAAACACTTGTGGGACCATCCCAAATGACAGAGCTCTACCACTCTTTAGC 1230
Qy |||||
1081 TGACCTTGAATAATGTGAGATTTCTAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 1140
Db |||||
1231 TGACCTTGAATAATGTGAGATTTCTAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 1290
Qy |||||
1141 GAAGGCCCTTTGCTGGATCTCTGAGCTGTGACGCTGTGATGCTGCTTGGACAGCA 1200
Db |||||
1291 GAAGGCCCTTTGCTGGATCTCTGAGCTGTGACGCTGTGATGCTGCTTGGACAGCA 1350
Qy |||||
1201 CAACCTCAAGCAAAATGACACCCATGGATATCTGAGATTTAATTTGTTGACCA 1260
Db |||||
1351 CAACCTCAAGCAAAATGACACCCATGGATATCTGAGATTTAATTTGTTGACCA 1410
Qy |||||
1261 TATTTATACCGCTGGAGCAAGAGCAACAAATTTGTTGTCACGCTCCCTCTCTGCTGGTGA 1320

1411 TATTTATACCGCTGGAGCAAGAGCAACAAATTTGGTCAACGTCCTCTCTGCTGGA 1470
Db |||||
1321 TATGCTCTGAACCTGGCTGCTGAATGTTATGATACGGGACGAACAGGAGGATCCCTGT 1380
Qy |||||
1471 TATGCTCTGAACCTGGCTGCTGAATGTTATGATACGGGACGAACAGGAGGATCCCTGT 1530
Db |||||
1381 CCTGCTTTTAAACTGCGCATCATTTCCCTGTGTAAAGCAATTTGGAAGCAAGTACAG 1440
Qy |||||
1531 CCTGCTTTTAAACTGCGCATCATTTCCCTGTGTAAAGCAATTTGGAAGCAAGTACAG 1590
Db |||||
1441 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACAGCGAGGCTGGGCT 1500
Qy |||||
1591 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACAGCGAGGCTGGGCT 1650
Db |||||
1501 C 1501
Db |||||
1651 C 1651

RESULT 2
US-09-845-416-10
; Sequence 10, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 3531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-10

Query Match 100.0%; Score 1501; DB 10; Length 3531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACATGCGCAAGACCTCCAAAGTGAATTTGAAGCTCACAGAGATGTTTATCAACACCTGGA 60
Db 1500 ACATGCGCAAGACCTCCAAAGTGAATTTGAAGCTCACAGAGATGTTTATCAACACCTGGA 1559
Qy 61 TGAACACAGCCAAATAATCTGAGATCCCTGGAAGTTCGATGATGCAAGTCTGTATCA 120
Db 1560 TGAACACAGCCAAATAATCTGAGATCCCTGGAAGTTCGATGATGCAAGTCTGTATCA 1619
Qy 121 AAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTTCGAAAAGTCTCTCAACAT 180
Db 1620 AAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTTCGAAAAGTCTCTCAACAT 1679
Qy 181 TAGGTCCCATTTGGAAGCCAGTTCGACCAAGTGGAAAGCGTCTGACCTTTCTCTGAGGA 240
Db 1680 TAGGTCCCATTTGGAAGCCAGTTCGACCAAGTGGAAAGCGTCTGACCTTTCTCTGAGGA 1739
Qy 241 ACTTCTGGTGTGGCTACAGCTGAAGATGATGAATTTAAGCCGCGCAGGACCTATTGGAGG 300
Db 1740 ACTTCTGGTGTGGCTACAGCTGAAGATGATGAATTTAAGCCGCGCAGGACCTATTGGAGG 1799
Qy 301 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 360
Db 1800 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 1859
Qy 361 AACTAAGAACCTGTAAATCATGAGTACTCTTGGAGACTGTACGAATTTCTGACAGAGCA 420
Db 1860 AACTAAGAACCTGTAAATCATGAGTACTCTTGGAGACTGTACGAATTTCTGACAGAGCA 1919

Qy	901	GCAGCTGCATGAAGACCCACAGGGAC	TTGGTCAGCATCTCAGCATCTTCTTTCACGTC	960
Db	2868	GCAGCTGCATGAAGCCACAGGGAC	TTGGTCAGCATCTCAGCATCTTCTTTCACGTC	2927
Qy	961	TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA	1020	
Db	2928	TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA	2987	
Qy	1021	CGAGACTCAAAACAACCTTGCTGGAGACCATCCCAAAATGACAGAGCTCTACAGTCTCTTTAGC	1080	
Db	2988	CGAGACTCAAAACAACCTTGCTGGAGACCATCCCAAAATGACAGAGCTCTACAGTCTCTTTAGC	3047	
Qy	1081	TGACCTGAAATTAATGTCCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA	1140	
Db	3048	TGACCTGAAATTAATGTCCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA	3107	
Qy	1141	GAAGGCCCTTTGCTTGAGTCTCTTGAGCCTGTGCAGCTGCATGTGATGTCCTTGACACAGCA	1200	
Db	3108	GAAGGCCCTTTGCTTGAGTCTCTTGAGCCTGTGCAGCTGCATGTGATGTCCTTGACACAGCA	3167	
Qy	1201	CAACCTCAAGCAAAATGACGAGCCATGGATATCCCTGCAGATTATTAATGTTTGACCCAC	1260	
Db	3168	CAACCTCAAGCAAAATGACGAGCCATGGATATCCCTGCAGATTATTAATGTTTGACCCAC	3227	
Qy	1261	TATTTATGACCGCTGGAGCAAGACACAAATTTGGTCAACGTCCTCTCTGCGTGGGA	1320	
Db	3228	TATTTATGACCGCTGGAGCAAGACACAAATTTGGTCAACGTCCTCTCTGCGTGGGA	3287	
Qy	1321	TATGTGTCGAACTGGCTGCTGTAATGTTTATGATACGGACGAAACAGGGAGGATCCGTGT	1380	
Db	3288	TATGTGTCGAACTGGCTGCTGTAATGTTTATGATACGGACGAAACAGGGAGGATCCGTGT	3347	
Qy	1381	CCGTGCTCTTTTAAAAACTGGCATATTTCCCTGTGTAAAGACATTTGGGAAGACAAGTACAG	1440	
Db	3348	CCGTGCTCTTTTAAAAACTGGCATATTTCCCTGTGTAAAGACATTTGGGAAGACAAGTACAG	3407	
Qy	1441	ATACCTTTTCACGCAAGTGGCAGTCTCAACAGGATTTTGTGACAGCGGAGGCTGGGCCT	1500	
Db	3408	ATACCTTTTCACGCAAGTGGCAGTCTCAACAGGATTTTGTGACAGCGGAGGCTGGGCCT	3467	
Qy	1501	C 1501		
Db	3468	C 3468		

```

RESULT 5
US-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US200301713.2A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 50/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-2

```

	Query Match	Best Local Similarity	Score 1501;	DB 10;	Length 4182;
	Matches 1501;	Conservative	0;	Mismatches	0;
				Indels	Gaps
QY	1	ACAATGGCAGAGACTCC	AAGTGAATTTGAAGCTCACACAGATGTTTATCACAACTGGA	60	
Db	2151	ACAATGGCAGAGACTCC	AAGTGAATTTGAAGCTCACACAGATGTTTATCACAACTGGA	2210	

Qy	61	TGAAAACAGCCAAAAAATCTCTGAGTACCCCTGGAAAGGTTCCGATGATCGACGTCCTGTTACA	120
Db	2211	TGAAAACAGCCAAAAAATCTCTGAGTACCCCTGGAAAGGTTCCGATGATCGACGTCCTGTTACA	2270
Qy	121	AAGACGTTTGGATAACAATGAACTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACAT	180
Db	2271	AAGACGTTTGGATAACAATGAACTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACAT	2330
Qy	181	TAGGTCCTATTTGGAGCCAGTTCTGACAGTGGAAAGCGTCTGCACTTTCTCTGCAAGGA	240
Db	2331	TAGGTCCTATTTGGAGCCAGTTCTGACCACTGGAAAGCGTCTGCACTTTCTCTGCAAGGA	2390
Qy	241	ACTTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCGCGGAGGACCTATTGTGAGG	300
Db	2391	ACTTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCGCGGAGGACCTATTGTGAGG	2450
Qy	301	CGACTTTCCAGCAGTTCCAGAAAGCAGAACGATGATCATATAGGCGCTTCAAGAGGGAAATTTGAA	360
Db	2451	CGACTTTCCAGCAGTTCCAGAAAGCAGAACGATGATCATATAGGCGCTTCAAGAGGGAAATTTGAA	2510
Qy	361	AACATAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGATATTTCTGCACAGAGCA	420
Db	2511	AACATAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGATATTTCTGCACAGAGCA	2570
Qy	421	GCCTTTGGAAAGGACTAGAGAAATCTTACAGGAGGCCAGAGAGCTGCTCTCTGAGGAGAG	480
Db	2571	GCCTTTGGAAAGGACTAGAGAAATCTTACAGGAGGCCAGAGAGCTGCTCTCTGAGGAGAG	2630
Qy	481	AGCCAGAAATGTCACTCGGCTTCTAGAAAGCAGGCTGAGAGGTCAATATCTGAGTGGGA	540
Db	2631	AGCCAGAAATGTCACTCGGCTTCTAGAAAGCAGGCTGAGAGGTCAATATCTGAGTGGGA	2690
Qy	541	AAATTCAACTGCACATCCGCTGACTCGCAGAGAGAAATAGATCAGACCCCTTGAAGAAGCT	600
Db	2691	AAATTCAACTGCACATCCGCTGACTCGCAGAGAGAAATAGATCAGACCCCTTGAAGAAGCT	2750
Qy	601	CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGAT	660
Db	2751	CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGAT	2810
Qy	661	CAAGGATCTGGCAGCGCGTGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACTTCGA	720
Db	2811	CAAGGATCTGGCAGCGCGTGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACTTCGA	2870
Qy	721	GAAAGTCAAGGCACCTTCGAGAGAGAAATTTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCAA	780
Db	2871	GAAAGTCAAGGCACCTTCGAGAGAGAAATTTGGCCCTCTGAAAGAGAACGCTGAGCCACGTCAA	2930
Qy	781	TGACCTTGTCTGCCAGCTTTACCACTTTGGGGATTCAGCTCTCACCGTATTAACCTCAGCAC	840
Db	2931	TGACCTTGTCTGCCAGCTTTACCACTTTGGGGATTCAGCTCTCACCGTATTAACCTCAGCAC	2990
Qy	841	TCTCGAGACCTGAAACCCAGATGGAGCTTCTCGAGGTGCGCTCTGAGGAGCCAGGTTCAG	900
Db	2991	TCTCGAGACCTGAAACCCAGATGGAGCTTCTCGAGGTGCGCTCTGAGGAGCCAGGTTCAG	3050
Qy	901	GCAGCTGCATGAAGCCCAACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGTC	960
Db	3051	GCAGCTGCATGAAGCCCAACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGTC	3110
Qy	961	TGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA	1020
Db	3111	TGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCA	3170
Qy	1021	CGAGACTCAAAACAACTTGTCTGGGACCATCTCCAAAATGACAGAGCTTACCAGTCTTTAGC	1080
Db	3171	CGAGACTCAAAACAACTTGTCTGGGACCATCTCCAAAATGACAGAGCTTACCAGTCTTTAGC	3230
Qy	1081	TGACCTGAAATATGTCTGAGTTCTCAGCTTATAGGACTTGCCATGAACTCCGAAAGACTGCA	1140
Db	3231	TGACCTGAAATATGTCTGAGTTCTCAGCTTATAGGACTTGCCATGAACTCCGAAAGACTGCA	3290

QY 1141 GAGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGTGTGATGCTTGGACCAGCA 1200
Db |
QY 3291 GAGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGTGTGATGCTTGGACCAGCA 3350
Db |
QY 1201 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGAGATTAATTAATGTTGACCA 1260
Db |
QY 3351 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGAGATTAATTAATGTTGACCA 3410
Db |
QY 1261 TATTTATGAGCGCTGAGCAAGAGCAACAATTTGGTCAAGTCCCTCTCTGCGTGA 1320
Db |
QY 3411 TATTTATGAGCGCTGAGCAAGAGCAACAATTTGGTCAAGTCCCTCTCTGCGTGA 3470
Db |
QY 1321 TATGTCGTGAATGCTGCTGAATGTTATGATACGGAGCAACAGGAGGATCGTGT 1380
Db |
QY 3471 TATGTCGTGAATGCTGCTGAATGTTATGATACGGAGCAACAGGAGGATCGTGT 3530
Db |
QY 1381 CTTGCTTTTAAACTGGGATCAATTCCTGCTGTAAGCACAATTTGGAAGACAAGTACAG 1440
Db |
QY 3531 CTTGCTTTTAAACTGGGATCAATTCCTGCTGTAAGCACAATTTGGAAGACAAGTACAG 3590
Db |
QY 1441 ATACCTTTTCAAGCAAGTGCAGTTCACAGGATTTTGTGACCGCAGGCTGGCCT 1500
Db |
QY 3591 ATACCTTTTCAAGCAAGTGCAGTTCACAGGATTTTGTGACCGCAGGCTGGCCT 3650
Db |
QY 1501 C 1501
Db 3651 C 3651

RESULT 6

US-09-845-416-30
; Sequence 30, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT FILING DATE: 2001-04-30
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 4498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-30

Query Match 100.0%; Score 1501; DB 10; Length 4498;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAATGCGAAGACCTCCAGGTGAATTCAGCTCACAGATGTTATCAACCTGGA 60
Db 2257 ACAATGCGAAGACCTCCAGGTGAATTCAGCTCACAGATGTTATCAACCTGGA 2316
QY 61 TGAACACGCAACCAAAATCCTGAGATCCCTGGAAGTTCGAGTATGACGCTCTGTACA 120
Db 2317 TGAACACGCAACCAAAATCCTGAGATCCCTGGAAGTTCGAGTATGACGCTCTGTACA 2376
QY 121 AGACCTTTGGTAAACATCACTCAAGTGGAGTGAACCTCGAAGAGTCTCTCAACAT 180
Db 2377 AGACCTTTGGTAAACATCACTCAAGTGGAGTGAACCTCGAAGAGTCTCTCAACAT 2436
QY 181 TAGGTCCTTATGGAAGCCAGTTCAGCAGTTCGAGGAGTTCGACCTTCTCTGCAGGA 240
Db 2437 TAGGTCCTTATGGAAGCCAGTTCAGCAGTTCGAGGAGTTCGACCTTCTCTGCAGGA 2496
QY 241 ACTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG 300
Db 2497 ACTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG 3556

QY 301 CGACTTTCAGCAGTTTCAGAAAGCAGAACGATGTATCATAGGSCCTTCAAGAGGGAATTGAA 360
Db 2557 CGACTTTCAGCAGTTTCAGAAAGCAGAACGATGTATCATAGGSCCTTCAAGAGGGAATTGAA 2616
QY 361 AACTAAAGAACTCTTAATCATGACTACTCTTGGAGCTGTAGATATTTCTGACAGAGCA 420
Db 2617 AACTAAAGAACTCTTAATCATGACTACTCTTGGAGCTGTAGATATTTCTGACAGAGCA 2676
QY 421 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCTCTGAGGAGAG 480
Db 2677 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCTCTGAGGAGAG 2736
QY 481 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATTAAGTGGGA 540
Db 2737 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATTAAGTGGGA 2796
QY 541 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGACT 600
Db 2797 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGACT 2856
QY 601 CCAGAACTTCAAGAGCCAGGATGAGCTGAGACTCAAGCTGGCCCAAGCTGAGGTGAT 660
Db 2857 CCAGAACTTCAAGAGCCAGGATGAGCTGAGACTCAAGCTGGCCCAAGCTGAGGTGAT 2916
QY 661 CAAGGATCTTGGCAGCCGCTGGCGATCTCTCAATTGACTCTCTCCCAAGTCACTCGA 720
Db 2917 CAAGGATCTTGGCAGCCGCTGGCGATCTCTCAATTGACTCTCTCCCAAGTCACTCGA 2976
QY 721 GAAAGTCAAGGCACTTCAGAGAGAAATTCGCTCTGAAAGAGAACTGAGCCACCTCAA 780
Db 2977 GAAAGTCAAGGCACTTCAGAGAGAAATTCGCTCTGAAAGAGAACTGAGCCACCTCAA 3036
QY 781 TGACCTTGTCTGCCAGCTTACACTTTGGGCATTCAGCTCTCACCTGTATTAACCTCAGC 840
Db 3037 TGACCTTGTCTGCCAGCTTACACTTTGGGCATTCAGCTCTCACCTGTATTAACCTCAGC 3096
QY 841 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAAGTGGCCGCTGAGACCCGAGTCA 900
Db 3097 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAAGTGGCCGCTGAGACCCGAGTCA 3156
QY 901 GCAGCTGCATGAAGCCACAGGACCTTTGGTTCAGATCTCAGACCTTTCTTCCAGCTC 960
Db 3157 GCAGCTGCATGAAGCCACAGGACCTTTGGTTCAGATCTCAGACCTTTCTTCCAGCTC 3216
QY 961 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCAAAACAAAGTCCCTTACTATCAACCA 1020
Db 3217 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCAAAACAAAGTCCCTTACTATCAACCA 3276
QY 1021 CGAGACTCAAAACAACTTGTGGGAGCCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 1080
Db 3277 CGAGACTCAAAACAACTTGTGGGAGCCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 3336
QY 1081 TGACCTGAATAAGTTCAGATTTCTAGCTATAGAGCTGCCATGAACTCCGAAGACTGCA 1140
Db 3337 TGACCTGAATAAGTTCAGATTTCTAGCTATAGAGCTGCCATGAACTCCGAAGACTGCA 3396
QY 1141 GAAGGCCCTTTGCTTGGATCTCTTGAACCTTGTGAGCTGTGAGTGTGATGCTTGGACAGCA 1200
Db 3397 GAAGGCCCTTTGCTTGGATCTCTTGAACCTTGTGAGCTGTGAGTGTGATGCTTGGACAGCA 3456
QY 1201 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGAGATTAATTAATGTTGACCA 1260
Db 3457 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGAGATTAATTAATGTTGACCA 3516
QY 1261 TATTTATGAGCGCTGAGCAAGAGCAACAATTTGGTCAAGTCCCTCTCTGCGTGA 1320
Db 3517 TATTTATGAGCGCTGAGCAAGAGCAACAATTTGGTCAAGTCCCTCTCTGCGTGA 3576
QY 1321 TATGTCGTGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCGTGT 1380
Db 3577 TATGTCGTGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCGTGT 3636

QY 1381 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAAGTACAG 1440
Db |
QY 3637 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAAGTACAG 3696
Db |
QY 1441 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGAGATTTTGTACACAGCGCAGGCTGGGCT 1500
Db |
QY 3697 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGAGATTTTGTACACAGCGCAGGCTGGGCT 3756
Db |
QY 1501 C 1501
Db 3757 C 3757

RESULT 7

US-09-845-416-29
; Sequence 29, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 29
; LENGTH: 4825
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-29

Query Match 100.0%; Score 1501; DB 10; Length 4825;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAATGGCAAGACCTCAAGGTGAAATTTGAAGCTCACAGATGTTTATCAACCTGGA 60
Db 2584 ACAATGGCAAGACCTCAAGGTGAAATTTGAAGCTCACAGATGTTTATCAACCTGGA 2643
QY 61 TGAAGACAGCAAAATCTGAGATCCCTGGAGGTTCCGATGATGCGAGTCTGTATCA 120
Db 2644 TGAAGACAGCAAAATCTGAGATCCCTGGAGGTTCCGATGATGCGAGTCTGTATCA 2703
QY 121 AAGACGTTTCGATAACATGAACTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACAT 180
Db 2704 AAGACGTTTCGATAACATGAACTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACAT 2763
QY 181 TAGGTCCATTTGGAGCCAGTTCTGACAGTGGAGCGTCTGCACCTTTCTCTGCAGGA 240
Db 2764 TAGGTCCATTTGGAGCCAGTTCTGACAGTGGAGCGTCTGCACCTTTCTCTGCAGGA 2823
QY 241 ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG 300
Db 2824 ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG 3063
QY 301 CGACTTTCAGAGCTTCAGAGCAGACGATGTACATAGGCGCTTCAGAGGGAATTTGAA 360
Db 2884 CGACTTTCAGAGCTTCAGAGCAGACGATGTACATAGGCGCTTCAGAGGGAATTTGAA 2943
QY 361 AACTTAAGAACCTTGTATCATGAGTACTCTTGAGACTGTACGAAATTTCTGCAGAGCA 420
Db 2944 AACTTAAGAACCTTGTATCATGAGTACTCTTGAGACTGTACGAAATTTCTGCAGAGCA 3003
QY 421 GCCTTTTGGAGGACTAGAGAACTTACAGAGAGCCAGAGAGCTGCTCCTCAGGAGAG 480
Db 3004 GCCTTTTGGAGGACTAGAGAACTTACAGAGAGCCAGAGAGCTGCTCCTCAGGAGAG 3063
QY 481 AGCCCAAGATCTACTCGGCTTCTACGAAGCAGGCTGAGAGGTCAATACTCAGTGGGA 540
Db 3064 AGCCCAAGATCTACTCGGCTTCTACGAAGCAGGCTGAGAGGTCAATACTCAGTGGGA 3123

RESULT 8

US-09-845-416-35
; Sequence 35, Application US/09845416

Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
FILE REFERENCE: DB1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 35
LENGTH: 4848
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-35

Query Match 100.0%; Score 1501; DB 10; Length 4848;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ACATGCGACGCTCAAGGTGAATTTGAAGTCTACACAGATGTTTATCAACACCTGGA	60
Db	2607	ACATGCGACGCTCAAGGTGAATTTGAAGTCTACACAGATGTTTATCAACACCTGGA <td>2666</td>	2666
QY	61	TGAACAGCCAAATAATCTGAGATCCCTGGAGGTTCCGATGATGCGAGTCTCTGTACA	120
Db	2667	TGAACAGCCAAATAATCTGAGATCCCTGGAGGTTCCGATGATGCGAGTCTCTGTACA	2726
QY	121	AAGCGTTTGGATAACATGAATTCAGTGGAGTGAATTCGGAAGGTTCCGATGATGCGAGTCTCTGTACA	180
Db	2727	AAGCGTTTGGATAACATGAATTCAGTGGAGTGAATTCGGAAGGTTCCGATGATGCGAGTCTCTGTACA	2786
QY	181	TAGTCCCATTTGGAAGCCAGTTCTGCAGAGTGGAGGTTCCGATGATGCGAGTCTCTGTACA	240
Db	2787	TAGTCCCATTTGGAAGCCAGTTCTGCAGAGTGGAGGTTCCGATGATGCGAGTCTCTGTACA	2846
QY	241	ACTCTGGTGGTCTACAGTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG	300
Db	2847	ACTCTGGTGGTCTACAGTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG	2906
QY	301	CGACTTCCAGCAGTTTCAGNAGCAGACGATGTACATAGGGCTTCAAGAGGGAATTGAA	360
Db	2907	CGACTTCCAGCAGTTTCAGNAGCAGACGATGTACATAGGGCTTCAAGAGGGAATTGAA	2966
QY	361	AACATAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA	420
Db	2967	AACATAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA	3026
QY	421	GCCTTTGGAAGGACTAGAGAACTTACAGAGAGCCAGAGAGTCCCTCTGAGAGAG	480
Db	3027	GCCTTTGGAAGGACTAGAGAACTTACAGAGAGCCAGAGAGTCCCTCTGAGAGAG	3086
QY	481	AGCCAGAGATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATCTGAGTGGGA	540
Db	3087	AGCCAGAGATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATCTGAGTGGGA	3146
QY	541	AAATTTGAACCTGCACTCCGCTGACTGGCAGAGAGAAATAGATGAGACCTTTGAAAGACT	600
Db	3147	AAATTTGAACCTGCACTCCGCTGACTGGCAGAGAGAAATAGATGAGACCTTTGAAAGACT	3206
QY	601	CCAGGAACCTCAGAGGCCAGGATGAGTGCACCTCAAGCTCGCCCAAGCTGAGGTAT	660
Db	3207	CCAGGAACCTCAGAGGCCAGGATGAGTGCACCTCAAGCTCGCCCAAGCTGAGGTAT	3266
QY	661	CAAGGATCTCTGCGACCCGCTGGCGATCTCTCTATTGATCTCTCCAAAGATCACTCGA	720
Db	3267	CAAGGATCTCTGCGACCCGCTGGCGATCTCTCTATTGATCTCTCCAAAGATCACTCGA	3326
QY	721	GAAGTCAAGGCACTTCGAGGAGAAATTCGGCTCTGAAAGAGAACTGAGCCAGCTCAA	780
Db	3327	GAAGTCAAGGCACTTCGAGGAGAAATTCGGCTCTGAAAGAGAACTGAGCCAGCTCAA	3386

RESULT 9

US-09-845-416-28
Sequence 28, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
FILE REFERENCE: DB1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 28
LENGTH: 4966
TYPE: DNA
ORGANISM: Homo sapiens

US-09-845-416-28

Query Match 100.0%; Score 1501; DB 10; Length 4966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAATGGCAAGACCTCCAAAGGTGAAATTTGAAGCTCACAGATGTTTATCAAACTGGA 60
DB 2725 ACAATGGCAAGACCTCCAAAGGTGAAATTTGAAGCTCACAGATGTTTATCAAACTGGA 2784

QY 61 TGAARACAGCCAAATAATCTCGATCCCTGGAAGGTTCCGATGATGAGTCTCTGTACA 120
DB 2785 TGAARACAGCCAAATAATCTCGATCCCTGGAAGGTTCCGATGATGAGTCTCTGTACA 2844

QY 121 AAGACGTTTGGTAAACATGAATCTCAAGTGGAGTGAATTCGGAATAAGTCTCTCAACAT 180
DB 2845 AAGACGTTTGGTAAACATGAATCTCAAGTGGAGTGAATTCGGAATAAGTCTCTCAACAT 2904

QY 181 TAGGTCCCAATTTGGAAGCCAGTCTGACAGGTGGAGCGTCTGCACCTTTCTCTCGAGA 240
DB 2905 TAGGTCCCAATTTGGAAGCCAGTCTGACAGGTGGAGCGTCTGCACCTTTCTCTCGAGA 2964

QY 241 ACTTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACTTATGGAGG 300
DB 2965 ACTTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACTTATGGAGG 3024

QY 301 CGACTTTCAGAGGTTTCAGAGCAGAAACGATGATGATGAGGCTTCAAGAGGGAATTGAA 360
DB 3025 CGACTTTCAGAGGTTTCAGAGCAGAAACGATGATGATGAGGCTTCAAGAGGGAATTGAA 3084

QY 361 AACTAAAGAACCTGTAAATCATGATGATCTCTGAGAGTGTACGAATATTTCTGACAGAGA 420
DB 3085 AACTAAAGAACCTGTAAATCATGATGATCTCTGAGAGTGTACGAATATTTCTGACAGAGA 3144

QY 421 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCGACAGAGCTCCCTCTGAGAGAG 480
DB 3145 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCGACAGAGCTCCCTCTGAGAGAG 3204

QY 481 AGCCAGATGTCACTCGGCTTCTAGAAAGAGGCTGAGGAGTCAATACAGTGGGA 540
DB 3205 AGCCAGATGTCACTCGGCTTCTAGAAAGAGGCTGAGGAGTCAATACAGTGGGA 3264

QY 541 AAAATTGAACTGCACTCCCTGACTGGCAGAGAAATAAGATGAGACCTTGAAGACT 600
DB 3265 AAAATTGAACTGCACTCCCTGACTGGCAGAGAAATAAGATGAGACCTTGAAGACT 3324

QY 601 CAGGAACTTCAAGAGCCACCGATGAGTGGACCTCAAGCTCGCCCAAGCTGAGTGAT 560
DB 3325 CAGGAACTTCAAGAGCCACCGATGAGTGGACCTCAAGCTCGCCCAAGCTGAGTGAT 3384

QY 661 CAAGGGATCTGGCAGCCGCTGGCGATCTCTCATTTGACTCTCTCAAGATCACCTCGA 720
DB 3385 CAAGGGATCTGGCAGCCGCTGGCGATCTCTCATTTGACTCTCTCAAGATCACCTCGA 3444

QY 721 GAAAGTCAAGGCACTTCGAGAGAAATTTGGCTCTGAAAGAGAAAGTGAGGCCACTCAA 780
DB 3445 GAAAGTCAAGGCACTTCGAGAGAAATTTGGCTCTGAAAGAGAAAGTGAGGCCACTCAA 3504

QY 781 TGACCTTCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGAC 840
DB 3505 TGACCTTCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGAC 3564

QY 841 TCTGGAAGACCTGAACACCCAGAGGACTTTGGTCCAGATCTCAGACCTTCTTCCAGCTC 900
DB 3565 TCTGGAAGACCTGAACACCCAGAGGACTTTGGTCCAGATCTCAGACCTTCTTCCAGCTC 3624

QY 901 GCAGCTGCATGAAGCCACAGGAGGACTTTGGTCCAGATCTCAGACCTTCTTCCAGCTC 960
DB 3625 GCAGCTGCATGAAGCCACAGGAGGACTTTGGTCCAGATCTCAGACCTTCTTCCAGCTC 3684

QY 961 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAAGTCCCTTACTATATCAACCA 1020
DB 3685 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAAGTCCCTTACTATATCAACCA 3744

RESULT 10

US-09-845-416-34

; Sequence 34, Application US/09845416

; Publication No. US20030171312A1

; GENERAL INFORMATION:

; APPLICANT: XIAO, XIAO

; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

; FILE REFERENCE: DE1142

; CURRENT APPLICATION NUMBER: US/09/845,416

; PRIOR FILING DATE: 2001-04-30

; PRIOR FILING DATE: 2000-04-28

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 34

; LENGTH: 4990

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-845-416-34

Query Match 100.0%; Score 1501; DB 10; Length 4990;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAATGGCAAGACCTCCAAAGGTGAAATTTGAAGCTCACAGATGTTTATCAAACTGGA 60
DB 2749 ACAATGGCAAGACCTCCAAAGGTGAAATTTGAAGCTCACAGATGTTTATCAAACTGGA 2808

QY 61 TGAARACAGCCAAATAATCTCGATCCCTGGAAGGTTCCGATGATGAGTCTCTGTACA 120
DB 2809 TGAARACAGCCAAATAATCTCGATCCCTGGAAGGTTCCGATGATGAGTCTCTGTACA 2868

QY 121 AAGACGTTTGGTAAACATGAATCTCAAGTGGAGTGAATTCGGAATAAGTCTCTCAACAT 180

Db 2869 AAGAGCTTTGGATTAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACAT 2928
Qy 181 TAGTCCCATTTTGAAGCCAGTCTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGA 240
Db 2929 TAGTCCCATTTTGAAGCCAGTCTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAGGA 2988
Qy 241 ACTTCTGGTGTGCTACAGCTGAAGATGATGAATTAAGCCGACGACCTATTGGAGG 300
Db 2989 ACTTCTGGTGTGCTACAGCTGAAGATGATGAATTAAGCCGACGACCTATTGGAGG 3048
Qy 301 CGACTTTCAGCAGTCTCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 360
Db 3049 CGACTTTCAGCAGTCTCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 3108
Qy 361 AACTAAGAACCTGTATATCATAGTACTCTTGAGACTGTGACGATATTTCTGACAGAGCA 420
Db 3109 AACTAAGAACCTGTATATCATAGTACTCTTGAGACTGTGACGATATTTCTGACAGAGCA 3168
Qy 421 GCCTTTGGAAGAGTACTAGAGAACTCTACAGAGAGCCAGAGAGTGCCTCCTGAGGAGAG 480
Db 3169 GCCTTTGGAAGAGTACTAGAGAACTCTACAGAGAGCCAGAGAGTGCCTCCTGAGGAGAG 3228
Qy 481 AGCCAGAAATGTCACTCGGCTTTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGA 540
Db 3229 AGCCAGAAATGTCACTCGGCTTTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGA 3288
Qy 541 AAAAAAGAACTCACTCGCTCACTGCGCAGAGAAAAATAGATGAGACCTTTGAAGACT 600
Db 3289 AAAAAAGAACTCACTCGCTCACTGCGCAGAGAAAAATAGATGAGACCTTTGAAGACT 3348
Qy 601 CCAGGAACCTTCAAGAGGCCACGAGTGAAGCTCAAGCTGCGCAAGCTGAGGTGAT 660
Db 3349 CCAGGAACCTTCAAGAGGCCACGAGTGAAGCTCAAGCTGCGCAAGCTGAGGTGAT 3408
Qy 661 CAAGGAGTCTGCGAGCCGTGGGCGATCTCTCAATGATCTCTCTCAAGATCACTGCA 720
Db 3409 CAAGGAGTCTGCGAGCCGTGGGCGATCTCTCAATGATCTCTCTCAAGATCACTGCA 3468
Qy 721 GAAAGTCAAGGCACTTCGAGAGAAATTTGGCTCTGAAAGAGAACTGAGCCAGCTGCA 780
Db 3469 GAAAGTCAAGGCACTTCGAGAGAAATTTGGCTCTGAAAGAGAACTGAGCCAGCTGCA 3528
Qy 781 TGACCTTGTGCGCAGCTTACCACTTTGGGCAATCAGCTCTCACCGTATAAOCCTCAGCAC 840
Db 3529 TGACCTTGTGCGCAGCTTACCACTTTGGGCAATCAGCTCTCACCGTATAAOCCTCAGCAC 3588
Qy 841 TCTGGAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCA 900
Db 3589 TCTGGAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCA 3648
Qy 901 GCAGCTGCATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCATCTTCTTCCACGTC 960
Db 3649 GCAGCTGCATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCATCTTCTTCCACGTC 3708
Qy 961 TGTCAGGGTCTCTGGAGAGGACCTCTCGGCAACAAAGTGCCTACTATATCAACA 1020
Db 3709 TGTCAGGGTCTCTGGAGAGGACCTCTCGGCAACAAAGTGCCTACTATATCAACA 3768
Qy 1021 CGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGCTCTTTAGC 1080
Db 3769 CGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGCTCTTTAGC 3828
Qy 1081 TGACCTGAATATGTGAGATCTCAGCTTATAGGACTGCGCAATGAACTCCGAGACTGCA 1140
Db 3829 TGACCTGAATATGTGAGATCTCAGCTTATAGGACTGCGCAATGAACTCCGAGACTGCA 3888
Qy 1141 GAAGGCCCTTTGCTTGGATCTCTTGAAGCTGTGAGTGTGATGCTTGGACAGCA 1200
Db 3889 GAAGGCCCTTTGCTTGGATCTCTTGAAGCTGTGAGTGTGATGCTTGGACAGCA 3948
Qy 1201 CAACCTCAAGAAAAATGACAGCCCATGGATATCTGAGATATTAATTTGTTGACCA 1260
Db 3949 CAACCTCAAGAAAAATGACAGCCCATGGATATCTGAGATATTAATTTGTTGACCA 4008

RESULT 11

US-09-845-416-36

; Sequence 36, Application US/09845416

; Publication No. US20030171312A1

; GENERAL INFORMATION:

; APPLICANT: XIAO, XIAO

; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

; FILE REFERENCE: DE1142

; CURRENT APPLICATION NUMBER: US/09/845,416

; CURRENT FILING DATE: 2001-04-30

; PRIOR FILING DATE: 2000-04-28

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 36

; LENGTH: 5060

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-845-416-36

Query Match 100.0%; Score 1501; DB 10; Length 5060;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAATGCCAAGACCTTCAAGGTGAAATTTGAAGTCAACAGATGTTTATCAACCTGGA 60
Db 2819 ACAATGCCAAGACCTTCAAGGTGAAATTTGAAGTCAACAGATGTTTATCAACCTGGA 2878
Qy 61 TGAACACAGCAAAATAATCTGAGATCCCTGGAAGGTTCGGATGATGCGAGTCTGTACA 120
Db 2879 TGAACACAGCAAAATAATCTGAGATCCCTGGAAGGTTCGGATGATGCGAGTCTGTACA 2938
Qy 121 AAGACGTTTGGATPAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACAT 180
Db 2939 AAGACGTTTGGATPAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACAT 2998
Qy 181 TAGGTCCCATTTGGAAGCCAGTTCACCAAGTGGAGCGTCTGCACCTTTCTCTGCAGGA 240
Db 2999 TAGGTCCCATTTGGAAGCCAGTTCACCAAGTGGAGCGTCTGCACCTTTCTCTGCAGGA 3058
Qy 241 ACTTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGACGACCTATTGGAGG 300
Db 3059 ACTTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGACGACCTATTGGAGG 3118
Qy 301 CGACTTTCAGCAGTTCAGAACGACGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 360
Db 3119 CGACTTTCAGCAGTTCAGAACGACGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 3178
Qy 361 AACTAAGAACCTTGAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 420

Db 3179 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGAGCTGTACGAATATTTCTGACAGAGCA 3238
Qy 421 GCCTTTGGAGGAGCTAGAGAACTCTTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAG 480
Db 3239 GCCTTTGGAGGAGCTAGAGAACTCTTACCAGGAGCCAGAGAGCTGCCTCTCTGAGGAGAG 3298
Qy 481 AGCCCAAGATGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGA 540
Db 3299 AGCCCAAGATGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGA 3358
Qy 541 AAAATTGAACCTGCATCTCGCTGACTGCGCAGAGAGAAAATAGATGAGACCCCTTGAAGACT 600
Db 3359 AAAATTGAACCTGCATCTCGCTGACTGCGCAGAGAGAAAATAGATGAGACCCCTTGAAGACT 3418
Qy 601 CCAGGAACCTTCAAGAGGCCACCGATGAGCTGGAACCTTCAAGCTGCGCAGAGCTGAGTGTAT 660
Db 3419 CCAGGAACCTTCAAGAGGCCACCGATGAGCTGGAACCTTCAAGCTGCGCAGAGCTGAGTGTAT 3478
Qy 661 CAAGGATCTGCGACGCGTGGCGCATCTCTCATTTGACTCTCTCCCAAGATCACTCGA 720
Db 3479 CAAGGATCTGCGACGCGTGGCGCATCTCTCATTTGACTCTCTCCCAAGATCACTCGA 3538
Qy 721 GAAAGTCAAGGCACTTCGAGGAGAAAATTCGGCTCTTGAAGAGAGAGCTGAGCCAGTCAA 780
Db 3539 GAAAGTCAAGGCACTTCGAGGAGAAAATTCGGCTCTTGAAGAGAGAGCTGAGCCAGTCAA 3598
Qy 781 TGAACCTTGTGCGCAGCTTACACCTTTGGGCAATTCAGCTCTCACCGTATTAACCTGAGCAC 840
Db 3599 TGAACCTTGTGCGCAGCTTACACCTTTGGGCAATTCAGCTCTCACCGTATTAACCTGAGCAC 3658
Qy 841 TCTGGAAGACCTCAACACAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGACCGAGTCAAG 900
Db 3659 TCTGGAAGACCTCAACACAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGACCGAGTCAAG 3718
Qy 901 GCAGCTGCATGGAAGCCACAGGAGCTTTGGTCAGCATCTCAGCACTTTCTTTTCCACGTC 960
Db 3719 GCAGCTGCATGGAAGCCACAGGAGCTTTGGTCAGCATCTCAGCACTTTCTTTTCCACGTC 3778
Qy 961 TGTCCAGGTCCTTGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 1020
Db 3779 TGTCCAGGTCCTTGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 3838
Qy 1021 CGAGACTCAAAACAACTTGTCTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTTAGC 1080
Db 3839 CGAGACTCAAAACAACTTGTCTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTTAGC 3898
Qy 1081 TGACCTGAATTAATGTGAGATCTCAGCTTATAGAGCTGCCATGGAATCCGAAGACTGCA 1140
Db 3899 TGACCTGAATTAATGTGAGATCTCAGCTTATAGAGCTGCCATGGAATCCGAAGACTGCA 3958
Qy 1141 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGTGCATGTGATGCTTGGACAGCA 1200
Db 3959 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGTGCATGTGATGCTTGGACAGCA 4018
Qy 1201 CAACCTCAAGCAAAATGACAGGCCATGGATATCTTGCAGATTAATTAATTTGTTGACCAC 1260
Db 4019 CAACCTCAAGCAAAATGACAGGCCATGGATATCTTGCAGATTAATTAATTTGTTGACCAC 4078
Qy 1261 TATTTATGAGCGCTGAGCAAGGAGCAACAAATTTGGTCAACGCTCCCTCTCTGCGTGA 1320
Db 4079 TATTTATGAGCGCTGAGCAAGGAGCAACAAATTTGGTCAACGCTCCCTCTCTGCGTGA 4138
Qy 1321 TATGTGTCTGAACTGGCTGCTGATTTATGATAGGGAACGAGGAGGAGTCCGTGT 1380
Db 4139 TATGTGTCTGAACTGGCTGCTGATTTATGATAGGGAACGAGGAGGAGTCCGTGT 4198
Qy 1381 CTTGTCTTTTAAACTGGCATATTTCCCTGTGTAAAGCAATTTGGAAGACAGTACAG 1440
Db 4199 CTTGTCTTTTAAACTGGCATATTTCCCTGTGTAAAGCAATTTGGAAGACAGTACAG 4258
Qy 1441 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGAGGCTGGGCT 1500
Db 4259 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGAGGCTGGGCT 4318

Qy 1501 C 1501
Db 4319 C 4319

RESULT 12

US-09-845-416-27
; Sequence 27, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 5149
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-27

Query Match 100.0%; Score 1501; DB 10; Length 5149;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAATGGCAAGACCTCCAAAGGTGAAATTTGAAGCTTCACACAGATGTTTATCACAACCTGGA 60
Db 2908 ACAATGGCAAGACCTCCAAAGGTGAAATTTGAAGCTTCACACAGATGTTTATCACAACCTGGA 2967
Qy 61 TGAATAACAGCAAAAAATTCCTGAGATCCCTGGAGGTTCCGATGATGAGTCCCTGTTTACA 120
Db 2968 TGAATAACAGCAAAAAATTCCTGAGATCCCTGGAGGTTCCGATGATGAGTCCCTGTTTACA 3027
Qy 121 AAGAGCTTTGATTAACATGAACCTTCAAGTGAGTGAACCTTCGGAAGAAAGTCTCTCAACAT 180
Db 3028 AAGAGCTTTGATTAACATGAACCTTCAAGTGAGTGAACCTTCGGAAGAAAGTCTCTCAACAT 3087
Qy 181 TAGGTCCCATTTGGAAGCCAGTTCTGACAGTGAAGCGTCTGCACCTTTCTCTGACGGA 240
Db 3088 TAGGTCCCATTTGGAAGCCAGTTCTGACAGTGAAGCGTCTGCACCTTTCTCTGACGGA 3147
Qy 241 ACTTCTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGAGGACCTTATTGGAGG 300
Db 3148 ACTTCTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGAGGACCTTATTGGAGG 3207
Qy 301 CGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTATAGGGCTTTCAAGAGGGAAATTGAA 360
Db 3208 CGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTATAGGGCTTTCAAGAGGGAAATTGAA 3267
Qy 361 AACTAAAGAACTGTATCATGAGTACTCTTGGAGCTGTAGCAATATTTCTGACAGAGCA 420
Db 3268 AACTAAAGAACTGTATCATGAGTACTCTTGGAGCTGTAGCAATATTTCTGACAGAGCA 3327
Qy 421 GCCTTTGGAAGGACTTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTCTGAGGAGAG 480
Db 3328 GCCTTTGGAAGGACTTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTCTGAGGAGAG 3387
Qy 481 AGCCCAAGATGTCACTCGGCTTCTACGAAACAGCGCTGAGAGGTCAATATCTGAGTGGGA 540
Db 3388 AGCCCAAGATGTCACTCGGCTTCTACGAAACAGCGCTGAGAGGTCAATATCTGAGTGGGA 3447
Qy 541 AAAATTGAACCTGCATCTCGCTGACTGCGCAGAGAAAATAGATGAGACCTTTGAAAGACT 600
Db 3448 AAAATTGAACCTGCATCTCGCTGACTGCGCAGAGAAAATAGATGAGACCTTTGAAAGACT 3507
Qy 601 CCAGGAACCTTCAAGAGGCCAGGATGAGCTGGAACCTTCAAGCTGGGCCAAGCTGAGGTGAT 660

Db 3508 CCAGGAACCTTCAAGAGGCCGAGGAGTGGAGCTTCAAGCTCGGCCAAGAGCTGAGTGAT 3567
 Qy 661 CAAGGATCTCTGGCAGCCCTGGGCGATCTCCCTCAITGACTCTCTCAAGATCAACCTCGA 720
 Db 3568 CAAGGATCTCTGGCAGCCCTGGGCGATCTCCCTCAITGACTCTCTCAAGATCAACCTCGA 3627
 Qy 721 GAAAGTCAAGGCACTTCGAGGAGAAATGGCCCTCTGAAAGAGAACTGAGCCACGTCAA 780
 Db 3628 GAAAGTCAAGGCACTTCGAGGAGAAATGGCCCTCTGAAAGAGAACTGAGCCACGTCAA 3687
 Qy 781 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTTAGCTCTCACCGTATTAACCTCGAC 840
 Db 3688 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTTAGCTCTCACCGTATTAACCTCGAC 3747
 Qy 841 TCTGGAAGACTGAACACCAAGAGTGGAGCTTCTGAGGTGGCGTTCGAGGACCGAGTCAG 900
 Db 3748 TCTGGAAGACTGAACACCAAGAGTGGAGCTTCTGAGGTGGCGTTCGAGGACCGAGTCAG 3807
 Qy 901 GCAGCTGCATGAAGCCGACAGGACTTTGGTCCAGCATCTCAGCAGCTTCTTCCACGTC 960
 Db 3808 GCAGCTGCATGAAGCCGACAGGACTTTGGTCCAGCATCTCAGCAGCTTCTTCCACGTC 3867
 Qy 961 TGTCCAGGCTCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCTTACTATCAACCA 1020
 Db 3868 TGTCCAGGCTCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCTTACTATCAACCA 3927
 Qy 1021 CGAGACTCAACAACTTGTGGGAGACCATCCCAAAATGACAGAGCTCTACAGCTTTTAGC 1080
 Db 3928 CGAGACTCAACAACTTGTGGGAGACCATCCCAAAATGACAGAGCTCTACAGCTTTTAGC 3987
 Qy 1081 TGACCTGAATAATGTCTCAGATTTCTCAGCTTATAGGACTGCGATGAAACTCCGAGAGTCA 1140
 Db 3988 TGACCTGAATAATGTCTCAGATTTCTCAGCTTATAGGACTGCGATGAAACTCCGAGAGTCA 4047
 Qy 1141 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGATGATGATGATGATGATGAT 1200
 Db 4048 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGATGATGATGATGATGATGAT 4107
 Qy 1201 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTAATTAATTTTTCAGCAC 1260
 Db 4108 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTAATTAATTTTTCAGCAC 4167
 Qy 1261 TATTTATGACCGCTGGAGCAAGAGCACAATTTGGTCAACGTCCTCTCTGGGTGA 1320
 Db 4168 TATTTATGACCGCTGGAGCAAGAGCACAATTTGGTCAACGTCCTCTCTGGGTGA 4227
 Qy 1321 TATGCTGCTGAAGTGGCTGCTGATTTTATGATAGGAGCAGCAAGGAGGATCCGTGT 1380
 Db 4228 TATGCTGCTGAAGTGGCTGCTGATTTTATGATAGGAGCAGCAAGGAGGATCCGTGT 4287
 Qy 1381 CTTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACAATTTGGAAGACAAGTACAG 1440
 Db 4288 CTTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACAATTTGGAAGACAAGTACAG 4347
 Qy 1441 ATACCTTTTCAAGCAAGTGGAGTCAACAGATTTTGTGACGAGGAGGCTGGGCT 1500
 Db 4348 ATACCTTTTCAAGCAAGTGGAGTCAACAGATTTTGTGACGAGGAGGCTGGGCT 4407
 Qy 1501 C 1501
 Db 4408 C 4408

RESULT 13
 US-10-149-736-40
 ; Sequence 40, Application US/10149736
 ; Publication No. US20030216332A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chamberlain, Jeffrey S.
 ; APPLICANT: Harper, Scott Q.
 ; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
 ; FILE REFERENCE: UM-06968
 ; CURRENT APPLICATION NUMBER: US/10/149,736

; CURRENT FILING DATE: 2002-06-17
 ; PRIOR APPLICATION NUMBER: PCT/US01/31126
 ; PRIOR FILING DATE: 2001-10-04
 ; PRIOR APPLICATION NUMBER: 60/238,848
 ; PRIOR FILING DATE: 2000-10-06
 ; NUMBER OF SEQ ID NOS: 96
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 40
 ; LENGTH: 5339
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic
 ; US-10-149-736-40

Query Match 100.0%; Score 1501; DB 16; Length 5339;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAATGGCAAGACCTCCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAACCTGGA 60
 Db 1693 ACAATGGCAAGACCTCCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAACCTGGA 1752
 Qy 61 TGAACACAGCCCAAAAAATCTGAGATCCCTGGAGGTTCGATGATGATGATGATGATGATGAT 120
 Db 1753 TGAACACAGCCCAAAAAATCTGAGATCCCTGGAGGTTCGATGATGATGATGATGATGATGAT 1812
 Qy 121 AAGACGTTTGGATAACATGAATTTCAAGTGGAGTGAATTTGGAAGAAAGTCTCTCAACAT 180
 Db 1813 AAGACGTTTGGATAACATGAATTTCAAGTGGAGTGAATTTGGAAGAAAGTCTCTCAACAT 1872
 Qy 181 TAGGTCCTATTTGGAGCCAGTCTGACAGTGGAGGCTCTGACGCTTCTCTGAGGAG 240
 Db 1873 TAGGTCCTATTTGGAGCCAGTCTGACAGTGGAGGCTCTGACGCTTCTCTGAGGAG 1932
 Qy 241 ACTTCTGCTGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGGACCTATTGGAGG 300
 Db 1933 ACTTCTGCTGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGGACCTATTGGAGG 1992
 Qy 301 CGATTTTCCAGCAGTTCAGAGCAGACAGATGTACATAGGCTTCAAGGGGAATTTGAA 360
 Db 1993 CGATTTTCCAGCAGTTCAGAGCAGACAGATGTACATAGGCTTCAAGGGGAATTTGAA 2052
 Qy 361 AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGCA 420
 Db 2053 AACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGCA 2112
 Qy 421 GCCTTTGGAAGGACTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTGAGAGAG 480
 Db 2113 GCCTTTGGAAGGACTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTGAGAGAG 2172
 Qy 481 AGCCCAAGATGTCACTCGGCTTCTAGAGAGAGGCTGAGGAGTCAATCTGAGTGGGA 540
 Db 2173 AGCCCAAGATGTCACTCGGCTTCTAGAGAGAGGCTGAGGAGTCAATCTGAGTGGGA 2232
 Qy 541 AAAATTGAACTCTGACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 600
 Db 2233 AAAATTGAACTCTGACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 2292
 Qy 501 CCAGGAACTTCAAGAGCCAGGATGAGCTGGACCTCAAGCTCGGCAAGCTGAGGTGAT 560
 Db 2293 CCAGGAACTTCAAGAGCCAGGATGAGCTGGACCTCAAGCTCGGCAAGCTGAGGTGAT 2352
 Qy 661 CAAGGATCTCTGGCAGCCCTGGGCGATCTCTCTCATTTGACTCTCTCCAAAGATCAACCTCGA 720
 Db 2353 CAAGGATCTCTGGCAGCCCTGGGCGATCTCTCTCATTTGACTCTCTCCAAAGATCAACCTCGA 2412
 Qy 721 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGGCTCTGAAAGAGAACTGAGCCACGTCAA 780
 Db 2413 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGGCTCTGAAAGAGAACTGAGCCACGTCAA 2472
 Qy 781 TGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTTAGCTCTCTCACCGTATTAACCTCAGCAC 840

Db 2473 TGACCTGTGCGCAGCTTACCACTTTGGCATTAGCTCTCACCGTATAAACCCTCAGCAC 2532
QY 841 TCTGGAAGCCTGAACACACAGATGGAAGCTTCTGCAGGTGGCGTGCAGGACCGAGTCA 900
Db 2533 TCTGGAAGCCTGAACACACAGATGGAAGCTTCTGCAGGTGGCGTGCAGGACCGAGTCA 2592
QY 901 GCAGCTGCATGAAGCCGACAGGACCTTGGTCCAGCATCTCAGCATTCTTTTCCAGTC 960
Db 2593 GCAGCTGCATGAAGCCGACAGGACCTTGGTCCAGCATCTCAGCATTCTTTTCCAGTC 2652
QY 961 TGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTACTATATCAACCA 1020
Db 2653 TGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTACTATATCAACCA 2712
QY 1021 CGAGACTCAAAACAACTTGTGGAGCCATCCAAAATGACAGAGCTCTACAGTCTTTAGC 1080
Db 2713 CGAGACTCAAAACAACTTGTGGAGCCATCCAAAATGACAGAGCTCTACAGTCTTTAGC 2772
QY 1081 TGACCTGAATAATGTGATGTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 1140
Db 2773 TGACCTGAATAATGTGATGTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 2832
QY 1141 GAAGGCCCTTGTGGATCTTGTAGCCTGTGAGCTGTGATGTGATGTGATGTGATGTGAT 1200
Db 2833 GAAGGCCCTTGTGGATCTTGTAGCCTGTGAGCTGTGATGTGATGTGATGTGATGTGAT 2892
QY 1201 CAACTCAAGCAAAATGACAGAGCCATGATATCTGCAAGTATTAATTTGTTGACCA 1260
Db 2893 CAACTCAAGCAAAATGACAGAGCCATGATATCTGCAAGTATTAATTTGTTGACCA 2952
QY 1261 TATTATGACCGCTGTGAGAGAGCAGCAAAATTTGGTCAACGCTCCCTCTCTGCGTGA 1320
Db 2953 TATTATGACCGCTGTGAGAGAGCAGCAAAATTTGGTCAACGCTCCCTCTCTGCGTGA 3012
QY 1321 TATGTGTGAACTGGTGTGATGTTTATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 3013 TATGTGTGAACTGGTGTGATGTTTATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAG 3072
QY 1381 CCTGTCTTTTAAACTGGCATCATTTCCCTGTCTAAAGACATTTTGAAGACACAGTACAG 1440
Db 3073 CCTGTCTTTTAAACTGGCATCATTTCCCTGTCTAAAGACATTTTGAAGACACAGTACAG 3132
QY 1441 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGTTTGTGACAGCGAGGCTGGGCT 1500
Db 3133 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGTTTGTGACAGCGAGGCTGGGCT 3192
QY 1501 C 1501
Db 3193 C 3193

RESULT 14
US-10-149-736-41
; Sequence 41, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 5462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Synthetic
US-10-149-736-41
Query Match 100.0%; Score 1501; DB 16; Length 5462;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAATGGCAAGACCTCCAAAGTGAATGAAGCTCACACAGATGTTTATCAACACCTGGA 60
Db 1816 ACAATGGCAAGACCTCCAAAGTGAATGAAGCTCACACAGATGTTTATCAACACCTGGA 1875
QY 61 TGAATAACAGCCAAATAATCCCTGGAAGTTCGATGATGATGATGATGATGATGATGATGAT 120
Db 1876 TGAATAACAGCCAAATAATCCCTGGAAGTTCGATGATGATGATGATGATGATGATGATGAT 1935
QY 121 AAGAGTTTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 180
Db 1936 AAGAGTTTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 1995
QY 181 TAGGTCCCATTTGGAAGCCAGTTCGACAGTGAAGCTGTGACACCTTTCTGTCAGGA 240
Db 1996 TAGGTCCCATTTGGAAGCCAGTTCGACAGTGAAGCTGTGACACCTTTCTGTCAGGA 2055
QY 241 ACTTCTGTGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db 2056 ACTTCTGTGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2115
QY 301 CGACTTTCAGCAGTTCAG 360
Db 2116 CGACTTTCAGCAGTTCAG 2175
QY 361 AACTTAAAGAACTGTAAATCATGAGTACTTCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 420
Db 2176 AACTTAAAGAACTGTAAATCATGAGTACTTCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 2235
QY 421 GCCTTTGGAAGAGACTAG 480
Db 2236 GCCTTTGGAAGAGACTAG 2295
QY 481 AGCCAGAGAGTGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATATGAGTGGGA 540
Db 2296 AGCCAGAGAGTGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATATGAGTGGGA 2355
QY 541 AAAATGTAACCTGCACTCCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 2356 AAAATGTAACCTGCACTCCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2415
QY 601 CCAGAGAGTTCAG 660
Db 2416 CCAGAGAGTTCAG 2475
QY 661 CAAGGAGATCTGGCAG 720
Db 2476 CAAGGAGATCTGGCAG 2535
QY 721 GAAAGTCAAGGAG 780
Db 2536 GAAAGTCAAGGAG 2595
QY 781 TGACCTTGTGCTGCAAGTTCACACTTTGGGCAATTCAGCTCTCAGCTATATAACCTCAGCAC 840
Db 2596 TGACCTTGTGCTGCAAGTTCACACTTTGGGCAATTCAGCTCTCAGCTATATAACCTCAGCAC 2655
QY 841 TGTGGAAG 900
Db 2656 TGTGGAAG 2715
QY 901 GCAGCTGATGAAG 960
Db 2716 GCAGCTGATGAAG 2775
QY 961 TGTCCAGGAGTCCCTGGGAG 1020

Db	4288	GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAAGTGTGATGCTTGGACCA	4347
QY	1201	CAACCTCAAGCAAAATGACCGCCATGGATATCCTGCAGATTATTAATTTGACCA	1260
Db	4348	CAACCTCAAGCAAAATGACCGCCATGGATATCCTGCAGATTATTAATTTGACCA	4407
QY	1261	TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGA	1320
Db	4408	TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGA	4467
QY	1321	TATGTGCTGAACTGCTGCTGAATGTTTATGATCGGACGAACAGGGAGGATCCGTGT	1380
Db	4468	TATGTGCTGAACTGCTGCTGAATGTTTATGATCGGACGAACAGGGAGGATCCGTGT	4527
QY	1381	CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAAGTACAG	1440
Db	4528	CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAAGTACAG	4587
QY	1441	ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCGCAGGCTGGGCT	1500
Db	4588	ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCGCAGGCTGGGCT	4647
QY	1501	C 1501	
Db	4648	C 4648	

Search completed: September 19, 2004, 15:14:15
Job time : 504.133 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 19, 2004, 02:48:31 ; Search time 2714.88 Seconds
(without alignments)
16510.166 Million cell updates/sec

Title: US-09-845-416-10_COPY_1500_3000
Perfect score: 1501
Sequence: 1 acaatggcgaagacctccaaag.....accagcgaggctgggcttc 1501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_esthum.*
4: em_esthum.*
5: em_esthum.*
6: em_esthum.*
7: em_esthum.*
8: em_esthum.*
9: gb_est1.*
10: gb_est2.*
11: gb_est2.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rnd.*
26: em_gss_phg.*
27: em_gss_vri.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	515.6	34.4	3753	11 AK081426	AK081426 Mus muscu
2	491.2	32.7	2874	29 AY408546	AY408546 Homo sapi
3	486.6	32.4	579	9 AL121550	AL121550 DKFZp762L
4	484	32.2	777	14 CD653550	CD653550 AGENCOURT

5	484	32.2	824	9	AL556247	AL556247
6	482.4	32.1	801	14	CB991394	CB991394 AGENCOURT
7	482.4	32.1	1098	13	EX365572	EX365572 AGENCOURT
8	482	32.1	797	14	CB960722	CB960722 AGENCOURT
9	477.6	31.8	2874	29	AY408546	AY408546 Mus muscu
10	475	31.6	834	12	BI729851	BI729851 603349511
11	473	31.5	620	13	BQ640063	BQ640063 he23G04.Y
12	462.8	30.8	508	14	CD701871	CD701871 EST18395
13	455.8	30.4	591	13	EX490860	EX490860 DKFZp886F
14	450.4	30.0	770	12	BG719710	BG719710 602690430
15	439	29.2	578	13	EX485574	EX485574 DKFZp686L
16	435	29.0	652	10	BB629984	BB629984 BB629984
17	435	29.0	1298	11	AK087829	AK087829 Mus muscu
18	435	29.0	2135	11	AK013510	AK013510 Mus muscu
19	435	29.0	4437	11	AK036936	AK036936 Mus muscu
20	433.4	28.9	1384	11	AK075809	AK075809 Mus muscu
21	430.2	28.7	704	10	BB610411	BB610411 BB610411
22	420.4	28.0	750	12	BI730168	BI730168 603349711
23	420.2	28.0	512	14	CF162938	CF162938 B0718G07-
24	417.4	27.8	663	12	BM488464	BM488464 pgm2n.pk0
25	412.4	27.5	854	9	AI196693	AI196693 ui53e10.Y
26	410.4	27.3	599	10	BB666688	BB666688 BB666688
27	410.2	27.3	3051	11	BC036095	BC036095 Homo sapi
28	407.2	27.1	554	14	CB613696	CB613696 AMGNNUC:N
29	406.6	27.1	665	13	BY742604	BY742604 BY742604
30	406.2	27.1	520	29	CG606443	CG606443 OST284316
31	404.4	26.9	493	14	CA888041	CA888041 B0142C06-
32	397.4	26.5	495	14	CA894775	CA894775 B0187G06-
33	389	25.9	644	13	BU313510	BU313510 603540290
34	385.6	25.7	650	13	BY714491	BY714491 BY714491
35	383.4	25.5	483	14	CA893902	CA893902 B0182B01-
36	381.8	25.4	423	9	AA460476	AA460476 zx61e10.r
37	381.8	25.4	513	29	CG605968	CG605968 OST283399
38	380.8	25.4	508	29	CG606947	CG606947 OST285335
39	379.4	25.3	717	14	CB527785	CB527785 UI-M-RY0-
40	369.8	24.6	488	29	CG597965	CG597965 OST261937
41	369.8	24.6	542	29	CG594848	CG594848 OST253426
42	369.2	24.6	502	9	AL871560	AL871560 AL871560
43	367.8	24.5	677	9	AL641565	AL641565 AL641565
44	366.2	24.4	622	9	AL796733	AL796733 AL796733
45	365.2	24.3	2874	29	AY408547	AY408547 Pan trogl

ALIGNMENTS

RESULT 1	AK081426	AK081426	3753 bp	linear	HTC 20-SEP-2003
LOCUS	AK081426	Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:CI30016x19	product:dystrophin related protein 2,		
DEFINITION	AK081426	full insert sequence.			
ACCESSION	AK081426	AK081426	GI:26349154		
VERSION	AK081426	HTC; CAP trapper.			
KEYWORDS	AK081426	Mus musculus (house mouse)			
SOURCE	AK081426	Mus musculus			
ORGANISM	AK081426	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
REFERENCE	AK081426	1			
AUTHORS	AK081426	Carninci, P. and Hayashizaki, Y.			
TITLE	AK081426	High-efficiency full-length cDNA cloning			
JOURNAL	AK081426	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	AK081426	99279253			
PUBMED	AK081426	10349636			
REFERENCE	AK081426	2			
AUTHORS	AK081426	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itch, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
TITLE	AK081426	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	AK081426	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	AK081426	20499374			
PUBMED	AK081426	11042159			

3
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tachiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Macumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wakahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer
JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE
20530513
PUBMED
11076861
4
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409, 685-690 (2001)
REFERENCE
AUTHORS
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL
Nature 420, 563-573 (2002)
REFERENCE
AUTHORS
6 (bases 1 to 3753)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, F., Saito, R., Saiton, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE
Direct Submission
JOURNAL
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
FEATURES
source
1. .3753
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB,C130016K19"
/db_xref="MGI:2413800"
/db_xref="taxon:10090"
/clone="C130016K19"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days embryo"
181_3054
/notes="unnamed protein product; dystrophin related protein 2 (MGJ|MGI:107432, GB|U43520, evidence: BLASTN, 99%, match=807)
putative"
/codon_start=1
/protein_id="BAC38217.1"
/db_xref="GI:26349155"

/translation="WQPLVMQGGCPYTLPRCHWEHAAADRFHSSSLRNTCPQPOVRAAV
TIPAPWDGADPCLSKPLNGTGTATGPLEPSAMNLCWNEIKKKHNLRLARLEAFSD
LSGKLQPLREILDMSQDEELSAQLPLOGDVALYQOEKETHAAWEEVKSGPYIS
SVLESAQAFISQPPPELESQSESKDTSRQRIQLSRFWQATVASELWEKLTAR
CVDHRIEHLTLEHLEIQAGMEELSSTLTQAGVRATWEPIDGLFIDSLPEHQAIK
LFKEEPFVKDGVKLVNDLAHQALISDVHLSMNSRALEGINIRWQLOQVSAERLQK
LQDARDFGSGQHFSLTSVQWEPAFSPENKVPYINHOAQTTCDHDPMTLSYOTL
ADLNKIFSAYRTAKLRVQKRLDLVLTLLTALBIFNEHDLQASAEHMDVVEVTHC
LTALYERLEBERGILVNVPLCVMSLWNLNVDSGRSRALSFKTGACUCGTEV
KEKLOYLFSOVANGSGCDDORHLGALLHEALQVPROLGEVAARFGSGNVSPVSRFRF
STKPVIEASQPLEWNLERSQMWLAVLHRTIAEQVKTKCSICROCPKGPYR
SLKQNVNIDICQTLGRASKNKLHPIMEYITPTTSSENDRFATLLNKRKRSQY
FSKQPGYLPVSVLESDCSETPASSPMLPHADTHRSIEHFASRLAESMNSQSPFN
DSLSPDDSIDQYLLHSSPITDRPFAQQAPCSMATSKGELSKXILAHLEENRI
LQELRLKWHHEAAEAPTVEGSAEATPDHNEELLAEARILQHKRLERMQIL
EDHNKLESOLQRLRELLQPPESDNGSAGSSLASPROSEGHPREKQTPDTE
VADDVGSKSODVSLCLEDIMEKLHRAFPVSRSSDVNTANTLAS"
3731..3736
/note="putative"
polyA_signal
3753
/note="putative"
polyA_site
ORIGIN
Query Match 34.4%; Score 515.6; DB 11; Length 3753;
Best Local Similarity 61.2%; Pred. No. 2.3e-141;
Matches 833; Conservative 0; Mismatches 529; Indels 0; Gaps 0;
136 CATGAATCTCTGTTGGAATGAATAAAGAGAGTCTCACAACTCCGCGCTCGCCTAGA 473
Db |||||
196 AGCCAGTCTGACAGTGGAGAGTCTGACCTTCTCTCCAGAGACTTCTGTGTGGCT 255
Qy |||||
474 GGCCTTCTGACACCTCAGTGGAAACTTTCAGTCCCGCTCCGAGAGATTATTGACTGGCT 533
Db |||||
256 ACAGCTCAAAAGATGATCAATTAAGCCGGCAGGCACCTATTGGAGGGGACCTTCCAGCAGT 315
Qy |||||
534 CAGCCAAAGAGTGGAGAGTGTTCAGCTCAGTCCCTTCGAGGGGATGTGCGCTGGT 593
Db |||||
316 TCAGAGCAGAACGATGTTACATAGGGCTTCAAGAGGGAATTGAAACTAAAGAACTGT 375
Qy |||||
594 ACAACAGAGAAAGAGACACATGACAGCTTTCATGGAAGAAAGTCAAGTCTAAGGGCCCTA 653
Db |||||
376 AATCATGAGTACTCTTTCAGAGCTGACGAATATTTCGACAGACAGCCTTTGGAAGACT 435
Qy |||||
654 CATCTCTGTGCTTGAATCCGCCAGGCTTTCGTCACGATCCATTGAAAGATT 713
Db |||||
436 AGAGAACTCTACAGGAGCCAGAGAGTGTCTCTGAGGAGAGAGCCAGAGATGTCAC 495
Qy |||||
714 AGAGGAGTCTCAATCTGAGAGTAAAGATACCTCCCCAGACAGCGGATTCAAACCTTAG 773
Db |||||
496 TCGGCTTTCAGAAAGCAGGCTGAGGAGGTCAATCTGAGTGGGAAATTTGACCTGCA 555
Qy |||||
774 TCGCTTTGTATGGAAGCAGGCAACAGTGGCCCTGAGTGTGGGAGAGCTGACGCCCG 833
Db |||||
556 CTCGCTGACTGGCAGAGAAATAATAGATGAGACCCCTTTGAAAGACTCCAGGAATTTCAAGA 615
Qy |||||
834 CTGTGTAGATCAGCATCGCCACATTGACATCTCTGGAACATCTATTGGAGATCCAAGG 893
Db |||||
616 GGCACAGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAGGGATCTTGGCA 875
Qy |||||
894 GGCAATGGAGAAATTGAGCAGTACTTTGACCCCAAGCAGAGGGAGTCCGAGCCACGTTGGGA 953
Db |||||
676 GCCCGTGGGGATCTCCTCATTTGACTCTCTCCAGATCACTCCAGAAAGTCAAGGCACT 735
Qy |||||
954 GCCATAGGAGATCTCTTTATTGATTCCTCCAGAGCATATCCAGCCATCAAGCTATT 1013
Db |||||
736 TCGAGGAGAAATTGGGCTCTGAAAGAGAACGTGAGCCAGCTCAATGACTTCTCTCGCCA 795
Qy |||||
1014 CAAGAAGAAATTTCTCTCTGAAAGATGGGGTGAAGTAGTGAATGATCTGCGCCACCA 1073
Db |||||
796 GCTTACCACCTTTGGGCAATTCAGTCTTCCCGTATACCTCAGCACTCTGGAACACCTGAA 855
Qy |||||

Db 1074 GCTTGCATCTCTGATGTGCACCTTGTCAATGGAGAAATTCAGGGCTCTGGAAACAGATCAA 1133
QY 856 CACAGATGGAAGCTTCTCAGAGTGGCCGTCGAGGACCGAGTCAAGCAGCTCATGAAGC 915
Db 1134 CATCCGGTGAACACAGCTCCAGGTGCAGTGGCTGAGAGGCTTAAGCAACTCCAGGATGC 1193
QY 916 CCACAGGGATTTGGTCCAGCATCTCAGACATCTTCTTCCACGCTCTGTCAGGGTCCCTG 975
Db 1194 CCACGGGACTTTGGGCTGGTTCACAGCACTTCCCTCTCCACTTGTCCAAATGTCCTG 1253
QY 976 GGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCCAGAGACTCAAAACAC 1035
Db 1254 GGAAGAGCATTTTCCCAATAAAGTCCCTACTACATCAACCCAGGCTCAGACCCAC 1313
QY 1036 TTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATATGT 1095
Db 1314 ATGCTGGGACCATCCAAAGATGAGTACTGATATACCAAGCCCTAGCGGATCTGAACACAT 1373
QY 1096 CAGATCTCAGCTTATAGGACTGCCATGAACCTCCGAAGCTGCAGAGGCGCTTTGCTT 1155
Db 1374 TAAGTCTCAGCTTACCGCACTGCCATGAAGTCCGAGTCCGAGTCCAGAGGCGCTTGGCT 1433
QY 1156 GGATCTCTGAGGCTGTGAGCTGCATGTGATGCTTGGACAGCAAACTCAAGCAAAA 1215
Db 1434 GGATCTGGTAACTTTAACTACAGCTCTGGAGATCTTCAATGAGCATGACTTGCAGGCCAG 1493
QY 1216 TGACCGCCCATGATATCTCGACATTAATTAATTTGTTGACCACTATTTATGACGCT 1275
Db 1494 TGAACATGATGATGTGGTGGAGTCAATCTACTGCTTGAAGTCTTGTATGAGCACT 1553
QY 1276 GGACGAAGACCAACAATTTGGTCAACGCTCCCTCTCTGCTGGATGATGTGTAAGCTG 1335
Db 1554 GGAGGAGGAAGAGGATCTCTGCTCAATGTGCCACTGTGTGAGATGAGGCTCAACTG 1613
QY 1336 GCTGCTCAATGTTATGATACGGAGCAAGACAGGAGGATCCGCTCTCTCTTTAAAC 1395
Db 1614 GCTCCTCAATGTTTGTATAGTGGTCCGAGTGAAGATGCGAGCATGTCCTTTAAGAC 1673
QY 1396 TGGCATCATTTCCCTGTGTAAGACACATTTGGAAGACAAAGTACAGATACCTTTCAAGCA 1455
Db 1674 TGGCATCGATGCTGTGGCAGGAGTGAAGAAAGTCAAGTCACTCTCTCAGCCA 1733
QY 1456 AGTGGCAAGTCAACAGGATTTGTGACGAGCCAGGCTGGG 1497
Db 1734 AGTAGCCAAATTCAGGACCGCAGTGTGATCAACGCCATCTCGG 1775

RESULT 2
AY408546 2874 bp DNA linear GSS 15-DEC-2003
LOCUS Homo sapiens DRP2 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY408546
VERSION AY408546.1 GI:39764517
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2874)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2874)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source 1..2874
organism="Homo sapiens"
mol_type="genomic DNA"
db_xref="taxon:9606"
<1..>2874
/gene="DRP2"
/locus_tag="HCM3260"

gene

ORIGIN

Query Match 32.7%; Score 491.2; DB 29; Length 2874;
Best Local Similarity 58.3%; Pred. No. 3.6e-134;
Matches 796; Conservative 0; Mismatches 570; Indels 0; Gaps 0;

QY 136 CATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGCTCTCAACATTAGTCCCAATTGGA 195
Db 234 CATGATCTGTGTGGATGAATAAAGAAAGTCTCACAACCTCCGCGCTCGCCTAGA 293
QY 196 AGCCAGTTCTCAACAGTGGAAAGCGTCTGCACCTTTCTCTGCAGAACTTCTGTGTGGCT 255
Db 294 GGCCTTCTCAGACACACAGTGGAAAGCTTCAGCTCCCTCTTCAAGAGATTATTGACTGGCT 353
QY 256 ACAGCTGAAGATGATGAATTAAGCCGCGAGCGACCTATTGGAGGCGACTTTCAGCAGT 315
Db 354 CAGCCAAAGATGAGGAGTTGTGAGTCACTGCCCTACAGGGGATGTGGCCCTGGT 413
QY 316 TCAGAAGCAGAACATGTACATAGGGCTTCAAGAGGGAATTGAAAATAAAGAACCTGT 375
Db 414 GCAACAGGAGAGGAGACACATGCGGCTTTATGGAAGAGTCAAGTCTCGGGGCCCTA 473
QY 376 AATCATGATGACTCTTGAGACTGTACGATATTTCTGACAGAGAGCTTTGGAGGACT 435
Db 474 CATCTATTCTGTGTGGAGTCAAGTCAAGCTTCTGTGCCAGACCCCAATTGAGGAGTT 533
QY 436 AGAGAAAATCTACACAGGAGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAAGATGTCA 495
Db 534 AGAGAGGCTCATCTGAGAGCAAGATACCTCCCGAAACAGCGGATCCAGATCTCAG 593
QY 496 TCGGCTTCTCAGAAAGAGGCTGAGGAGTCAATCTAGTGGGAAAAAATTGAACCTGCA 555
Db 594 CCGCTTTGTATGGAAGCAGGCGAGCGTGGCCAGTGAACCTGTGGAGAGATTGACAGCCCG 653
QY 556 CTCCGCTCAGCTGGCAGAGAAAATAGATGAGACCTTGAAGACTCCAGGAACTTCAAGA 615
Db 654 CTGTGTGACAGACACCTCATTTGAGCGGACTCTGGAGAGCTCTTGGAGATTCAGG 713
QY 616 GGGCCAGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGTGTATCAAGGATCCTGGCA 675
Db 714 GGCATGAGGAACTAAGCACTACTCTGAGCCAAGCTGAGGAGTCCGAGCCACTTGGGA 773
QY 676 GCGCGTGGCGATCTCTCATTTGACTCTCTCCAGATCACTCCGAGAAAGTCAAGGCACT 735
Db 774 GCGCATTTGGGATCTCTTATTTGATTTACTCCAGAGCACTCCAGGCTATTAGCTGTT 833
QY 736 TCGAGGAAAAATTGGCCTCTCAAAAGAGACGCTGAGCCACGCTCAATGACCTTGTCTGCCA 795
Db 834 CAAAGAAATTTCTCCCCCATGAAGATGAGTAAAGTTGTGATGATCTTGGCCACCA 893
QY 796 GCTTACCACTTTGGGCATTCAGCTCTCAACCTATAACCTCAGCACTCTGGAAGACCTGAA 855
Db 894 ACTTGCCATTTCTGATGTGCACCTTGTCAATGGAGAAATTCAGGCGCTTGGAAACAGATCAA 953
QY 856 CACCAAGATGGAAGCTTCTGCAAGTGGCGCTCGAGACCGAGTCAAGGAGCTGCAATGAAGC 915
Db 954 GCTCCGATGGAACCAACTACAGGCGTCAATTAGTAGAGGCTTAAGCACTCCAGGATGC 1013
QY 916 CCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGCTGTGTCAGGCTCCCTG 975

DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC c lone distribution information
 can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDAM513 row: a column: 23
 High quality sequence stop: 715.
 Location/Qualifiers

FEATURES

1..777
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30426742"
 /tissue_type="Embryonic Stem cells"
 /cell_line="WA01"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
 /notes="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI;
 This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01 cell line Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLN3, SSEA-1, TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with Trizol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199]) Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGATCTTCTAGATCGGAGCGCCGCTTTTITTTTITTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Loxe-linker L1-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The average insert size is about 3.6kb."

ORIGIN

Query Match 32.2%; Score 484; DB 14; Length 777;
 Best Local Similarity 100.0%; Pred. No. 2e-132;
 Matches 484; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1018 CCACGAGACTCAAAACACTTCTGGGACCATCCCAAATGACAGAGCTCTACAGTCTTT 1077
 DB 137 CCACGAGACTCAAAACACTTCTGGGACCATCCCAAATGACAGAGCTCTACAGTCTTT 196

QY 1078 AGCTGACCTGAATAATGTCAAGTTCTCAGCTTTATAGGACTGCCATGAATCCGAAGACT 1137
 DB 197 AGCTGACCTGAATAATGTCAAGTTCTCAGCTTTATAGGACTGCCATGAATCCGAAGACT 256

QY 1138 GCAGAGGCGCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGTGATGTGATGCCCTTGACCA 1197
 DB 257 GCAGAGGCGCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGTGATGTGATGCCCTTGACCA 316

QY 1198 GCACAACTCAAGCAAAATGACAGCCCATGGATATCCTGACATTATTAAATTTGTTGAC 1257
 DB 317 GCACAACTCAAGCAAAATGACAGCCCATGGATATCCTGACATTATTAAATTTGTTGAC 376

QY 1258 CACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAAGCTCCCTCTCTCGGT 1317
 DB 377 CACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAAGCTCCCTCTCTCGGT 436

QY 1318 GGATATGTCTGTGAACCTGGCTGCTGAATGTTATGATACGGGACGACAGGAGATCCG 1377
 DB 437 GGATATGTCTGTGAACCTGGCTGCTGAATGTTATGATACGGGACGACAGGAGATCCG 496

QY 1378 TGTCTCTGTCTTTTAAAACTTGGCATCTTCCCTGTGTAAAGCACATTTTGAAGACAAGTA 1437
 DB 497 TGTCTCTGTCTTTTAAAACTTGGCATCTTCCCTGTGTAAAGCACATTTTGAAGACAAGTA 556

QY 1438 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGCGCGCTGGG 1497
 DB 557 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGCGCGCTGGG 616

QY 1498 CCTC 1501
 DB 617 CCTC 620

RESULT 5

AL556247
 LOCUS AL556247 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 DEFINITION CDNA clone CS0DK001YB17 5-PRIME, mRNA sequence.
 ACCESSION AL556247
 VERSION AL556247.2 GI:31278051
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:12898746.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqres@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 1955.r
 Contact: Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DK001CA09Q1.

FEATURES

source
 1..824
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DK001YB17"
 /cell_type="HELA CELLS COT 25-NORMALIZED"
 /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 32.2%; Score 484; DB 9; Length 824;
 Best Local Similarity 100.0%; Pred. No. 2.1e-132;
 Matches 484; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1018 CCACGAGACTCAAAACACTTCTGGGACCATCCCAAATGACAGAGCTCTACAGTCTTT 1077
 DB 195 CCACGAGACTCAAAACACTTCTGGGACCATCCCAAATGACAGAGCTCTACAGTCTTT 254

QY 1078 AGCTGACCTGAATAATGTCAAGTTCTCAGCTTTATAGGACTGCCATGAATCCGAAGACT 1137
 DB 255 AGCTGACCTGAATAATGTCAAGTTCTCAGCTTTATAGGACTGCCATGAATCCGAAGACT 314

QY 1138 GCAGAGGCGCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGTGATGTGATGCTTGGACCA 1197

```

Db 315 GCAGAGGCCCTTGTCTTGGATCTCTTGGCTGTGATGCTGATGCTTGGACCA 374
QY 1198 GCACAACTCAAGCAAAATGACAGCCATGGATATCTCTGAGATTAATTAATTTGAC 1257
Db 375 GCACAACTCAAGCAAAATGACAGCCATGGATATCTCTGAGATTAATTAATTTGAC 434
QY 1258 CACTATTATGACCGCTGGAGCAAGACCAAAATTTGGTCAAGTCTCTGCGT 1317
Db 435 CACTATTATGACCGCTGGAGCAAGACCAAAATTTGGTCAAGTCTCTGCGT 494
QY 1318 GGATATGCTGTAACCTGGCTGCTGAATGTTATGATACGGGACGACAGGAGGATCCG 1377
Db 495 GGATATGCTGTAACCTGGCTGCTGAATGTTATGATACGGGACGACAGGAGGATCCG 554
QY 1378 TGTCTGCTCTTTAAACTGGCATCTTCCCTGTGTAAAGACATTTGGAACAAGTA 1437
Db 555 TGTCTGCTCTTTAAACTGGCATCTTCCCTGTGTAAAGACATTTGGAACAAGTA 614
QY 1438 CAGATACCTTTCAAGCAAGTGGCAAGTTCAGCAGGATTTGTGACAGCGAGGCTGGG 1497
Db 615 CAGATACCTTTCAAGCAAGTGGCAAGTTCAGCAGGATTTGTGACAGCGAGGCTGGG 674
QY 1498 CCCC 1501
Db 675 CCTC 678

RESULT 6
CB991394
LOCUS
DEFINITION AGENCOURT 13627932 NIH_MGC_148 Homo sapiens cDNA clone
IMAGE:30336570 5', mRNA sequence.
ACCESSION CB991394
VERSION CB991394.1 GI:30285818
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 801)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM360 row: d column: 19
High quality sequence stop: 621.
Location/Qualifiers
1. .801
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue="IMAGE:30336570"
/lab_host="DH10B TONa"
/clone_lib="NIH_MGC_148"
/notes="Organ: Placenta; Vector: pBluescriptR; Site: 1;
all-XhoI; Site 2: BamHI; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTGTCTTTTGTGTTT-3', size-selected for average insert
size 2.3 kb and normalized to ROT 5. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NIH/NHGRI,
National Institutes of Health). Note: this is a NIH_MGC

```

```

Library."
Query Match 32.1%; Score 482.4; DB 14; Length 801;
Best Local Similarity 99.8%; Pred. No. 6.1e-12;
Matches 483; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1018 CCACGAGACTCAAAACAACCTTCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTT 1077
Db 140 CCACGAGACTCAAAACAACCTTCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTT 199
QY 1078 AGCTGACCTGAATTAATGTACAGTCTTATAGACTGCCATGAAATCTCGAAGACT 1137
Db 200 AGCTGACCTGAATTAATGTACAGTCTTATAGACTGCCATGAAATCTCGAAGACT 259
QY 1138 GCAGAAGCCCTTTGCTTGGATCTCTTGGCCTGTGATGCTGATGCTGCTGGACCA 1197
Db 260 GCAGAAGCCCTTTGCTTGGATCTCTTGGCCTGTGATGCTGATGCTTGGACCA 319
QY 1198 GCACAACTCAAGCAAAATGACAGCCATGGATATCTCTGAGATTAATTAATTTGAC 1257
Db 320 GCACAACTCAAGCAAAATGACAGCCATGGATATCTCTGAGATTAATTAATTTGAC 379
QY 1258 CACTATTATGACCGCTGGAGCAAGACCAAAATTTGGTCAAGTCTCTCTCGT 1317
Db 380 CACTATTATGACCGCTGGAGCAAGACCAAAATTTGGTCAAGTCTCTCTCGT 439
QY 1318 GGATATGCTGTAACCTGGCTGCTGAATGTTATGATACGGGACGACAGGAGGATCCG 1377
Db 440 GGATATGCTGTAACCTGGCTGCTGAATGTTATGATACGGGACGACAGGAGGATCCG 499
QY 1378 TGTCTGCTCTTTAAACTGGCATCTTCCCTGTGTAAAGACATTTGGAACAAGTA 1437
Db 500 TGTCTGCTCTTTAAACTGGCATCTTCCCTGTGTAAAGACATTTGGAACAAGTA 559
QY 1438 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAGCAGGATTTGTGACAGCGAGGCTGGG 1497
Db 560 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAGCAGGATTTGTGACAGCGAGGCTGGG 619
QY 1498 CCTC 1501
Db 620 CCTC 623

RESULT 7
BX365572
LOCUS
DEFINITION BX365572 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DK001YB21 5-PRIME, mRNA sequence.
ACCESSION BX365572
VERSION BX365572.1 GI:30366927
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1098)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact: Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS1AK001ZB11QP1.
FEATURES
source
1. .1098
/organism="Homo sapiens"
/mol_type="mRNA"

```

```
/db_xref="taxon:9606"
/clone="GSDK001YB21"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

```
Query Match 32.1%; Score 482.4; DB 13; Length 1098;
Best Local Similarity 99.8%; Pred. No. 7.7e-132;
Matches 483; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1018 CCACGAGACTCAAAACACTTCTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTT 1077
DB 204 CCACGAGACTCAAAACACTTCTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTT 263

QY 1078 AGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAATCCGAAGACT 1137
DB 264 AGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAATCCGAAGACT 323

QY 1138 GCAGAAGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGCTGCTGATGTCCTTTGACCA 1197
DB 324 GCAGAAGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGCTGCTGATGTCCTTTGACCA 383

QY 1198 GCACAACCTCAAGAAAATGACAGGCCATGGATATCTCTGAGATTATTAATTTTGAC 1257
DB 384 GCACAACCTCAAGAAAATGACAGGCCATGGATATCTCTGAGATTATTAATTTTGAC 443

QY 1258 CACTATTATGACCGCTTGAGAGAGAGACACAAATTTGGTCAAGCTCCCTCTCTGCGT 1317
DB 444 CACTATTATGACCGCTTGAGAGAGAGACACAAATTTGGTCAAGCTCCCTCTCTGCGT 503

QY 1318 GGATATGTCTGAACCTGCTGCTGAATGTTTATGATACGGAGCAACAGGGAGGATCCG 1377
DB 504 GGATATGTCTGAACCTGCTGCTGAATGTTTATGATACGGAGCAACAGGGAGGATCCG 563

QY 1378 TGTCTGTCTTTTAAACTGGCATCTTCCCTGTGTAAAGACATTTGGAAGACAAGTA 1437
DB 564 TGTCTGTCTTTTAAACTGGCATCTTCCCTGTGTAAAGACATTTGGAAGACAAGTA 623

QY 1438 CAGATACCTTTTCAAGCAAGTGGCAAGTTTGTGACCAAGGATTTTGTGACCGCAGGCTGG 1497
DB 624 CAGATACCTTTTCAAGCAAGTGGCAAGTTTGTGACCAAGGATTTTGTGACCGCAGGCTGG 683

QY 1498 CCTC 1501
DB 684 CCTC 687
```

```
RESULT 8
CB960722 797 bp mRNA linear EST 29-APR-2003
LOCUS
DEFINITION
AGENCOURT 13761995 NIH MGC 147 Homo sapiens cDNA clone
IMAGE:30344479 5', mRNA sequence.
ACCESSION
CB960722
VERSION
CB960722.1 GI:30216839
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 797)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hanson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: NDM380 row: n column: 08
High quality sequence stop: 620.
Location/Qualifiers
1. 797
source

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30344479"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_147"
/notes="Organ: placenta; Vector: pBluescriptR; Site: 1:  
ali-xhoI; Site 2: BamH; Oligo-dr primed using primer  
5'-TTTATTTTATTTTATTTT-3', size-selected for average  
insert size 2.3 kb and normalized to ROT 5. This is a  
primary library enriched for full-length clones and  
constructed using the Cap-trapper method (Carninci, in  
preparation). Library constructed by M. Brownstein  
(NIH/NHGRI, National Institutes of Health). Note: This is  
a NIH_MGC library."
```

ORIGIN

```
Query Match 32.1%; Score 482; DB 14; Length 797;
Best Local Similarity 100.0%; Pred. No. 8.1e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1018 CCACGAGACTCAAAACACTTCTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTT 1077
DB 168 CCACGAGACTCAAAACACTTCTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTT 227

QY 1078 AGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAATCCGAAGACT 1137
DB 228 AGCTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAATCCGAAGACT 287

QY 1138 GCAGAAGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGCTCATGTGATGCTTTGGACCA 1197
DB 288 GCAGAAGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGCTCATGTGATGCTTTGGACCA 347

QY 1198 GCACAACCTCAAGCAAAATGACAGGCCATGGATATCTCTGAGATTATTAATTTTGAC 1257
DB 348 GCACAACCTCAAGCAAAATGACAGGCCATGGATATCTCTGAGATTATTAATTTTGAC 407

QY 1258 CACTATTATGACCGCTTGAGAGAGAGACACAAATTTGGTCAAGCTCCCTCTCTGCGT 1317
DB 408 CACTATTATGACCGCTTGAGAGAGAGACACAAATTTGGTCAAGCTCCCTCTCTGCGT 467

QY 1318 GGATATGTCTGAACCTGCTGCTGAATGTTTATGATACGGAGCAACAGGGAGGATCCG 1377
DB 468 GGATATGTCTGAACCTGCTGCTGAATGTTTATGATACGGAGCAACAGGGAGGATCCG 527

QY 1378 TGTCTGTCTTTTAAACTGGCATCTTCCCTGTGTAAAGACATTTGGAAGACAAGTA 1437
DB 528 TGTCTGTCTTTTAAACTGGCATCTTCCCTGTGTAAAGACATTTGGAAGACAAGTA 587

QY 1438 CAGATACCTTTTCAAGCAAGTGGCAAGTTTGTGACCAAGGATTTTGTGACCGCAGGCTGG 1497
DB 588 CAGATACCTTTTCAAGCAAGTGGCAAGTTTGTGACCAAGGATTTTGTGACCGCAGGCTGG 647

QY 1498 CC 1499
DB 648 CC 649
```

```
RESULT 9
AY408548 2874 bp DNA linear GSS 15-DEC-2003
LOCUS
DEFINITION
Mus musculus DRP2 gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.
```

AY408548
 AY408548.1 GI:39764519
 GSS.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2874)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
 Adams, M.D. and Cargill, M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Science 302 (5652), 1960-1963 (2003)
 14671302
 2 (bases 1 to 2874)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
 Adams, M.D. and Cargill, M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 Location/Qualifiers
 source
 1..2874
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..>2874
 /gene="DRP2"
 /locus_tag="HCM3260"
 ORIGIN
 Query Match 31.8%; Score 477.6; DB 29; Length 2874;
 Best Local Similarity 57.7%; Pred. No. 4.2e-130;
 Matches 786; Conservative 0; Mismatches 576; Indels 0; Gaps 0;
 QY 136 CATGAATCTCAAGTGGAGTCAACTCGGAAAAGTCTCTCAACATTAGTGGCCCAATTGGA 195
 DB 234 CATGAATCTGTGTGGAAATGAAATAAAGAAAGTCTCAACCTCGCGCTGCGCTAGA 293
 QY 196 AGCCAGTCTTCAACAGTGGAGCGTCTGCAACCTTTCTGCAAGAACTTCTGTGTGGCT 255
 DB 294 GGCTTCTCAGACCTCAGTGGAAACTTCAGTCCCGCTCCGAGAGATTATTGACTGGCT 353
 QY 256 ACAGCTGAAGATGATGAATTAAGCCGCGAGCACTATTGGAGGCGACTTTCCAGCAGT 315
 DB 354 CAGCCAAAAGGATGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGGAGTGTGAGGAGTGTG 413
 QY 316 TCAGAACAGACGATGATACATAGGCGCTTCAAGAGGGAATGAAACATAAGAACCTGT 375
 DB 414 ACACAGGAGAGAGACATCCAGCTTTCATGGAAGAGTCAAGTCTAAGGCGCCCTA 473
 QY 376 AATCATGAGTACTTGTGAGACTGTACGAATATTTCTGACAGACGACCTTTTGAAGGACT 435
 DB 474 CATCTCTCTGTGCTTGAATCCGCCAGGCTTTTCTGTCCAGCATCCATTGGAAGAAAT 533
 QY 436 AGAGAAACTCTACAGAGCCAGAGAGTGGCTCTGAGGAGAGAGCCAGCATGTCAC 495
 DB 534 AGAGGAGTCTCAATCTGAGAGTAAAGATACCTCCCCAGACGCGAATTCAAAACCTTAG 593
 QY 496 TCGGCTTCTACGAAACAGGCTGAGGAGTCAATCTAGTGGGAAAAATTGAACCTGCA 555
 DB 594 TCGCTTTGTATGGAAGCAGCAACAGTGGCAGTGTGAGTGTGAGGAGCTGACAGCCG 653
 QY 556 CTGCGTGTGCTGACAGAGAAATAGATGAGACCTTGAAGACTCCAGGACTTCAAGA 615
 DB 654 CTGTGTAGATCAGCATCGCCACATTGAGCATCTCTGGAACATCTATTGGAGATCCAAAG 713
 QY 616 GGCCAGGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCA 675

DB 714 GGCAATGGAGGAATTTGAGCAGTACTTTGACCAAGCAGGAGGAGTCCGAGCAGCTGGGA 773
 QY 676 GCCCGTGGCGATCTCCTCATTTGACTCTCTCAAGATCAGCTCGAGAAAGTCAAGGCACT 735
 DB 774 GCCCATAGAGATCTCTTTATTTGATTCCTCCAGAGCATATCCAGGCATCAAGCTATT 833
 QY 736 TGGAGGAGAAATGGCGCTCTGAAAGAGAACGTGAGCACCGTCAATGACCTTGTCCGCCA 795
 DB 834 CAAGAAGAATTTCTCTCTGTGAAAGATGGGTGAAGTAGTGAATGATCTGGCCACCA 893
 QY 796 GCTTACCACTTTGGGCATTTGAGCTCTCACCGTATAAAGTCTGAGCAGCTTGGAGACCTGAA 855
 DB 894 GCTTGCCTATCTCTGATGTCACCTTTCAATGAGAAATTCAGGGCTCTGGAACAGATCAA 953
 QY 856 CACCAGATGGAAGCTTTCTGAGGTGGCGCTGAGAGCCGAGTCCAGGAGTGTGATGAGC 915
 DB 954 CATCCGTTGGAACACAGCTCCAGGTGTGAGTGTGAGAGGCTTAAGCAACTCCAGGATGC 1013
 QY 916 CCAGAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTCTCAGAGGTCCCTG 975
 DB 1014 CACCGGGACTTTGGGCGCTGGGTCCAGCACTTCTCTCCACTTCTCTCCAAGTTCCTG 1073
 QY 976 GGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAACAAC 1035
 DB 1074 GGAAAGAGCAATTTCTCCCAATAGTCCCTACTACATCAANNNNNNNNNNNNNNNN 1133
 QY 1036 TTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGCTTTTGTAGCTGACCTGATATGT 1095
 DB 1134 NNN 1193
 QY 1096 CAGATTCTCAGCTTATAGGCTGCAATGAACTCCGAGAGCTGAGAGGCGCTTTTCTGCTT 1155
 DB 1194 TAAGTTCAGCTTACCGCACTGCCATGAGCTCCGAGAGTCCAGAGGCTCCAGGCGCT 1253
 QY 1156 GGATCTCTTGAGCCTGTGAGTGCATGTGATGCTGGAGTGGAGTGGAGTGGAGTGGAGTGG 1215
 DB 1254 GGAATGTGTAATTTAACTACAGCTCTGGAGATCTTCAATGAGCATGACTTGCAGGCCAG 1313
 QY 1216 TGACAGGCGCATGATCTCTGAGATTTGATGATTTGATGATTTGATGATTTGATGATTTG 1275
 DB 1314 TGAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1373
 QY 1276 GGAGCAAGAGCACAACTTTGTCACAGCTCTCTCTGCTGGATGATGATGATGATGATGAT 1335
 DB 1374 GGAGGAGAAAGAGGCATCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1433
 QY 1336 GCTGCTGAATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1395
 DB 1434 GCTCCTCAATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1493
 QY 1396 TGGCATCATTTCTGCTGTAAGCACATTTGGAAGACAACTTACAGATACCTTTTCAAGCA 1455
 DB 1494 TGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1553
 QY 1456 AGTGGCAAGTTCACAGGATTTTGTGACGAGCGAGCTGGG 1497
 DB 1554 AGTAGCAATTCAGGACGAGTGTGATCAACGCCATCTCGG 1595
 RESULT 10
 BI729851
 LOCUS 603349511F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357162 5',
 DEFINITION mRNA sequence.
 ACCESSION BI729851
 VERSION BI729851.1 GI:15706864
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 834)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM11908 row: e column: 03
High quality sequence stop: 796.
Location/Qualifiers
1. 834
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5357162"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 31.6%; Score 475; DB 12; Length 834;
Best Local Similarity 91.0%; Pred. No. 1e-129;
Matches 527; Conservative 0; Mismatches 50; Indels 2; Gaps 2;
QY 1 ACAATGCCAGACCTCCAGGTGAATTAAGTCTACAGATGTTTATCACAACTGCA 60
Db 230 ACCATGCCAGACCTCCAGGAGAAATGAACCTACAGATATCTATCAAACTTCA 289
QY 61 TGAACACAGCAAAATCTCGATCCCTGGAGGTTCCGATGATGCGTCTGTTACA 120
Db 290 TGAATGGCCAAAATCTCGATCCCTGGAGGTTCCGATGATGCGTCTGTTACA 349
QY 121 AAGACGTTTGATTAACATGAATCTCAAGTGGAGTGAATTCGGAAGATCTTCAACAT 180
Db 350 AAGACGTTTGATTAACATGAATCTCAAGTGGAGTGAATTCGGAAGATCTTCAACAT 409
QY 181 TAGTCCCATTTGGAACCGATTTGACCGATGGAAGCTCTGACCTTTCTTCGAGGA 240
Db 410 TAGTCCCATTTGGAACCGATTTGACCGATGGAAGCTTTGATCTTTCTTCGAGGA 469
QY 241 ACTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGCACCTATTGGAGG 300
Db 470 ACTTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGCACCTATTGGAGG 529
QY 301 CGATTTCCAGCAGTTTCAAGAGCAGACGATGTACATAGGGCCCTTCAAGAGGAATTGAA 360
Db 530 TGATTTCCAGCAGTTTCAAGAGCAGACGATGTACATAGGGCCCTTCAAGAGGAATTGAA 589
QY 361 AACTAAAGACCTCTAATCATGACTACTTTGAGACTGTACGAATATTTCTGACAGACA 420
Db 590 AACTAAAGACCTCTAATCATGACTACTTTGAGACTGTACGAATATTTCTGACAGACA 649
QY 421 GCCTTTGGAAGGACTAGAGAACTC-TACCAGGAGCCAGAGAGCTCCCTTCCTGAGAGA 479
Db 650 GCCTTTGGAAGGACTAGAGAACTC-TACCAGGAGCCAGAGAGCTCCCTTCCTGAGAAA 709
QY 480 GAGCCAGAAATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGCTCAATCACTCAGTGGG 539
Db 710 GAGCTCAGAAATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGCTCAATCACTCAGTGGG 579
QY 540 AAAAATTGAACTCGCACTC-CGCTGACTGGCAGAGAAA 577
Db 770 ACCAATTGAACTCGCTCAGCTGATTTGGCAGAGAAA 808

RESULT 11

BQ640063
LOCUS

DEFINITION
he23g04.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he

ACCESSION
BQ640063

VERSION
BQ640063.1 GI:21764522

KEYWORDS
EST.

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE
1 (bases 1 to 620)

JOURNAL
Wistow, G., Bernstein, S.L., Wyatt, M.K., Ray, S., Behal, A.,

MEDLINE
Touchman, J.W., Bouffard, G., Smith, D., and Peterson, K.

COMMENT
Expressed sequence tag analysis of human retina for the NEIBank

Project: Retbindin, an abundant, novel retinal cDNA and alternative

splicing of other retina-preferred gene transcripts

Mol. Vis. 8 (4), 196-204 (2002)

22103461

12107411

Contact: Wistow G

Section on Molecular Structure and Function

National Eye Institute

6/331, NIH, Bethesda, MD 20892-2740, USA

Tel: 301 402 3452

Fax: 301 496 0078

Email: graeme@helix.nih.gov

Plate: 23 row: g column: 04

Seq primer: M13RPI reverse primer (ABI).

Location/Qualifiers

1. 620

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="he23g04"

/tissue_type="Retina"

/dev_stage="Adult"

/lab_host="EMDH10B"

/clone_lib="Human Retina cDNA (Un-normalized,

unamplified): hd/he"

BQ640063 620 bp mRNA linear EST 15-JUL-2002

he23g04.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he

Homo sapiens cDNA clone he23g04 5', mRNA sequence.

BQ640063

Homo sapiens (human)

EST.

1 GI:21764522

31.5%; Score 473; DB 13; Length 620;

Best Local Similarity 100.0%; Pred. No. 3.3e-129;

Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1029 AAACAACTTGTGGGACCATCCAAATGACAGAGCTTACAGCTTTAGTGCCTGA 1088

Db 1 AAACAACTTGTGGGACCATCCAAATGACAGAGCTTACAGCTTTAGTGCCTGA 60

QY 1089 ATAAATGTGAGTTCTCAGCTTATAGGACTGCCATGAAATCCGAAAGACTGCAGAGGCC 1148

Db 61 ATAAATGTGAGTTCTCAGCTTATAGGACTGCCATGAAATCCGAAAGACTGCAGAGGCC 120

QY 1149 TTTCCTTGGATCTCTTGAGCCCTGTGAGTGTGATGTGCTTGGACAGACACCTCA 1208

ORIGIN

Query Match 31.5%; Score 473; DB 13; Length 620;
Best Local Similarity 100.0%; Pred. No. 3.3e-129;
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1029 AAACAACTTGTGGGACCATCCAAATGACAGAGCTTACAGCTTTAGTGCCTGA 1088
Db 1 AAACAACTTGTGGGACCATCCAAATGACAGAGCTTACAGCTTTAGTGCCTGA 60
QY 1089 ATAAATGTGAGTTCTCAGCTTATAGGACTGCCATGAAATCCGAAAGACTGCAGAGGCC 1148
Db 61 ATAAATGTGAGTTCTCAGCTTATAGGACTGCCATGAAATCCGAAAGACTGCAGAGGCC 120
QY 1149 TTTCCTTGGATCTCTTGAGCCCTGTGAGTGTGATGTGCTTGGACAGACACCTCA 1208

```

Db      121  TTGCTTGATCTCTTGAGCCTGTGAGTGCATGTGATCCCTTGGACGACCAACCTCA 180
QY      1209  AGCAAAATGACGAGCCCATGATATCTCTGACAGATTATTAATTTGTTGACCACTATTTATG 1268
Db      181  AGCAAAATGACGAGCCCATGATATCTCTGACAGATTATTAATTTGTTGACCACTATTTATG 240
QY      1269  ACCGCTGGAGCAAGAGACCAACAAATTTGTCACAGTCCCTCTCTGCGTGATATGTGTC 1328
Db      241  ACCGCTGGAGCAAGAGACCAACAAATTTGTCACAGTCCCTCTCTGCGTGATATGTGTC 300
QY      1329  TGAAGTGCCTCTGAATGTTTATCATCGGACGACAGGAGGATCCGTGTCCTGTCCT 1388
Db      301  TGAAGTGCCTCTGAATGTTTATCATCGGACGACAGGAGGATCCGTGTCCTGTCCT 360
QY      1389  TTAAGTGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACCTTT 1448
Db      361  TTAAGTGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACCTTT 420
QY      1449  TCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACGAGCGAGCTGGGCTC 1501
Db      421  TCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACGAGCGAGCTGGGCTC 473

RESULT 12
CD701871
LOCUS   CD701871
DEFINITION EST18395 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION CD701871
VERSION   CD701871.1 GI:32232501
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 508)
AUTHORS  Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and Zeng, Y.-X.
TITLE    Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL  Unpublished (2003)
COMMENT  Contact: Yixin Zeng
        Cancer Center
        Sun Yat-sen University
        651 Dongfeng Road East, Guangzhou 510060, China
        Tel: 86-1380-9770-743
        Fax: 86-20-8775-4506
        Email: yxzeng@zsums.edu.cn.
FEATURES             source
         1..508
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="normal nasopharynx"
            /clone_lib="human nasopharynx"
            /notes="ESTs generated from a normal nasopharynx cDNA
            library from southern Chinese"

ORIGIN
Query Match          30.8%; Score 462.8; DB 14; Length 508;
Best Local Similarity 99.6%; Pred. No. 3.1e-126;
Matches 464; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1036  TTGCTGGACCATCCAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAATGT 1095
Db      16   TTGCTGGACCATCCAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAATGT 75
QY      1096  CAGATTCTCAGCTTATAGAGTCCCATGAATCTCCGAGACTGCAGAGCCCTTTGCTT 1155
Db      76   CAGATTCTCAGCTTATAGAGTGCCTGAATGAATCTCCGAGACTGCAGAGCCCTTTGCTT 135
QY      1156  GGATCTCTTTGAGCTGTGAGTGCATGTGATGCCCTTGGACGACCAACCTCAAGCAAAA 1215
Db      136  GGATCTCTTTGAGCTGTGAGTGCATGTGATGCCCTTGGACGACCAACCTCAAGCAAAA 195

```

```

QY      1216  TGACCAGCCCATGGATATCCCTGACAGATTATTAATTTGTTTACCACCTATTTATGACCGCT 1275
Db      196  TGACCAGCCCATGGATATCCCTGACAGATTATTAATTTGTTTACCACCTATTTATGACCGCT 255
QY      1276  GGAGCAAGAGACCAACAAATTTGGTCAACGTCCTCTCTGCGTGATATGTGTAACG 1335
Db      256  GGAGCAAGAGACCAACAAATTTGGTCAACGTCCTCTCTGCGTGATATGTGTAACG 315
QY      1336  GCTGCTGAATGTTTATGATACGGACGACAGGAGGATCCGTGTCCTGTCCTTTTAAAC 1395
Db      316  GCTGCTGAATGTTTATGATACGGACGACAGGAGGATCCGTGTCCTGTCCTTTTAAAC 375
QY      1396  TGGCATCATTTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCA 1455
Db      376  TGGCATCATTTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCA 435
QY      1456  AGTGGCAAGTTCACAGGATTTTGTGACGAGCGAGCTGGGCTC 1501
Db      436  AGTGGCAAGTTCACAGGATTTTGTGACGAGCGAGCTGGGCTC 481

RESULT 13
BX490860
LOCUS   BX490860
DEFINITION DKFZp686F1392_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
ACCESSION DKFZp686F1392.5, mRNA sequence.
VERSION   BX490860
KEYWORDS  BX490860.1 GI:32000785
SOURCE   EST.
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 591)
AUTHORS  Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.
TITLE    EST (Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., et al.)
JOURNAL  Unpublished (2003)
COMMENT  Contact: MIPS
        MIPS
        Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
        This is the 5' sequence of the clone insert
        Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
        Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
        Sequenced by BHPZ (Biomedical Research Center at the Heinrich-
        Heine-University, Dueseldorf/Germany) within the cDNA sequencing
        consortium of the German Genome Project. No s1 sequence available.
        This clone (DKFZp686F1392) is available at the RZPD in Berlin.
        Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
        Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             source
         1..591
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DKFZp686F1392"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="686 (synonym: hlcc3)"
            /notes="Vector: pTriplex2; Site_1: SfiI; Site_2: SfiIB;
            cDNA-collection"

ORIGIN
Query Match          30.4%; Score 455.8; DB 13; Length 591;
Best Local Similarity 99.3%; Pred. No. 4.3e-124;
Matches 457; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1018  CCACGAGACTCAAAACACTTGTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTT 1077
Db      132  CCACGAGACTCAAAACACTTGTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTT 191
QY      1078  AGCTGACCTGAATGTCAGATCTTCAGCTTATAGACTGCATGAACCTCGAAGAT 1137

```

Db	192	AGTGACCTGAATAATGTTCAGATTCTCAGCTTATAGGACTGCCATGAATCCGAGACT	251
Qy	1138	GCAGAGGCCCTTGTCTGGATCTCTTGGAGCTGTCAGCTGCATGTGATGCTTGGACCA	1197
Db	252	GCAGAGGCCCTTGTCTGGATCTCTTGGAGCTGTCAGCTGCATGTGATGCTTGGACCA	311
Qy	1198	GCACAACTCAAGCAAAATGACAGCCCATGATATCTTGCAGATTAATTTGTTGAC	1257
Db	312	GCACAACTCAAGCAAAATGACAGCCCATGATATCTTGCAGATTAATTTGTTGAC	371
Qy	1258	CACATTTATGACCCCTGGAGCAGAGCAGACACAAATTTGGTCAACGTCCTCTCTCGGT	1317
Db	372	CACATTTATGACCCCTGGAGCAGAGCAGACACAAATTTGGTCAACGTCCTCTCTCGGT	431
Qy	1318	GGATATGCTGCTCAACTGGCTGCTGAAATTTTATGATACGGGACGAAACAGGAGGATCCG	1377
Db	432	GGATATGCTGCTCAACTGGCTGCTGAAATTTTATGATACGGGACGAAACAGGAGGATCCG	491
Qy	1378	TGTCCTGCTTTTAAACCTGGATCATTTCCCTGTGTAAAGCACATTTGGAGACAAGTA	1437
Db	492	TGTCCTGCTTTTAAACCTGGATCATTTCCCTGTGTAAAGCACATTTGGAGACAAGTA	551
Qy	1438	CAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTT	1477
Db	552	CAGATACCTTTTCAAGCANGTGGCAAGTTTCAACAGGATTT	591
RESULT 14	BG719710	602690430F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4822807 5',	
LOCUS	BG719710	770 bp mRNA linear EST 08-MAY-2001	
DEFINITION	BG719710	mRNA sequence.	
ACCESSION	BG719710	GI:13998897	
VERSION	BG719710	1	
KEYWORDS	BG719710	EST.	
SOURCE	BG719710	Homo sapiens (human)	
ORGANISM	BG719710	Homo sapiens	
REFERENCE	BG719710	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	BG719710	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE	BG719710	1 (Bases 1 to 770)	
JOURNAL	BG719710	NIH-MGC http://mgc.nci.nih.gov/.	
COMMENT	BG719710	National Institutes of Health, Mammalian Gene Collection (MGC)	
UNPUBLISHED	BG719710	Unpublished (1999)	
CONTACT	BG719710	Contact: Robert Strausberg, Ph.D.	
EMAIL	BG719710	Email: ccapbs-@mail.nih.gov	
TISSUE	BG719710	Tissue Procurement: Miklos Palkovits, M.D., Ph.D.	
CDNA	BG719710	cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki	
LIBRARY	BG719710	Toshiyuki and Piero Carninci (RIKEN)	
CDNA	BG719710	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
SEQUENCING	BG719710	DNA Sequencing by: Incyte Genomics, Inc.	
DISTRIBUTION	BG719710	Clone distribution: MGC clone distribution information can be	
FOUND	BG719710	found through the I.M.A.G.E. Consortium/LLNL at:	
HTTP	BG719710	http://image.llnl.gov	
PLATE	BG719710	Plate: LIAM10731 row: 1 column: 08	
QUALITY	BG719710	High quality sequence stop: 767.	
LOCATION	BG719710	Location/Qualifiers	
FEATURES	BG719710	1..770	
SOURCE	BG719710	/organism="Homo sapiens"	
	BG719710	/mol_type="mRNA"	
	BG719710	/db_xref="taxon:9606"	
	BG719710	/clone="IMAGE:4822807"	
	BG719710	/lab_host="DH10B"	
	BG719710	/clone_lib="NIH_MGC_97"	
	BG719710	/note="Organ: testis; Vector: pBluescriptR (modified	
	BG719710	pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI	
	BG719710	(gtcgag); Cligo-dT primed using primer	
	BG719710	5'-TTTTTTTTTTTTTTTNN-3', size-selected for average	
	BG719710	insert size 2.2 kb and normalized to Rof 5. this is a	
	BG719710	primary library enriched for full-length clones and	
	BG719710	constructed using the Cap-trapper method (Carninci,	
	BG719710	preparation). Library constructed by M. Brownstein	
	BG719710	(NIH/NHGRI, National Institutes of Health). Note: this is	
	BG719710	NIH_MGC library"	

```
source      1. .578
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp586L01247"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="896 (synonym: hlcc3)"
/note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN
Query Match      29.2%; Score 439; DB 13; Length 578;
Best Local Similarity 99.5%; Pred. No. 4.3e-119;
Matches 439; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1018 CCAGGAGACTCAAACTTGGCTGGGACCATCCAAAATGACAGAGCTCTACCCAGTCTTT 1077
Db 138 CCAGGAGACTCAAACTTGGCTGGGACCATCCAAAATGACAGAGCTCTACCCAGTCTTT 197
Qy 1078 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 1137
Db 198 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 257
Qy 1138 GCAGAAGGCCCTTTGGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCCTTGGACCA 1197
Db 258 GCAGAAGGCCCTTTGGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCCTTGGACCA 317
Qy 1198 GCACAACTTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAAATTGTTGAC 1257
Db 318 GCACAACTTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAAATTGTTGAC 377
Qy 1258 CACTATTATGACCGCTGGAGCAAGACACACAAATTTGGTCAAGCTCCCTCTCTGCGT 1317
Db 378 CACTATTATGACCGCTGGAGCAAGACACAAATTTGGTCAAGCTCCCTCTCTGCGT 437
Qy 1318 GGATATGTGTCTGAAGTGGCTGCTGAATGTTATGATACGGGACGAAACAGGGAGGATCG 1377
Db 438 GGATATGTGTCTGAAGTGGCTGCTGAATGTTATGATACGGGACGAAACAGGGAGGATCG 497
Qy 1378 TGTCCCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAGCACATTTTGGAAAGCAAGTA 1437
Db 498 TGTCCCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAGCACATTTTGGAAAGCAAGTA 557
Qy 1438 CAGATACCTTTTCAAGCAAGT 1458
Db 558 CAGATACCTTTTCAAGCAAGT 578
```

Search completed: September 19, 2004, 11:02:19
Job time : 2719.88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2004, 18:59:56 ; Search time 3974.06 Seconds
(without alignments)
16370.617 Million cell updates/sec

Title: US-09-845-416-12_COPY_1500_3000

Perfect score: 1501

Sequence: 1 agaagactctgagacaagaac.....ttctgcatgattctatccaa 1501

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.nam.*
37: em.htg.vit.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred.No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1328.8	88.5	5339	6	AX538620 Sequence
2	1328.8	88.5	5462	6	AX538621 Sequence
3	1328.8	88.5	5952	6	AR304538 Sequence
4	1328.8	88.5	5952	6	AX114289 Sequence
5	1328.8	88.5	8689	6	AX538622 Sequence
6	1328.8	88.5	11443	6	AX538624 Sequence
7	1328.8	88.5	12057	6	AX538627 Sequence
8	1328.8	88.5	12446	9	HSDMDR
9	1328.8	88.5	13957	6	AX409637 Sequence
10	1328.8	88.5	13957	6	AX538581 Sequence
11	1328.8	88.5	13957	9	HUMDYS
12	1317.8	87.8	13977	6	AR220819 Sequence
13	1185.4	79.0	13887	4	AF070485
14	1149.6	76.6	13815	6	AX306153 Sequence
15	1149.6	76.6	13815	6	AX538582 Sequence
16	1149.6	76.6	13815	10	MUSDYSA
17	1149.6	76.6	19307	6	AR093392 Sequence
18	1149.6	76.6	19307	6	AR142592 Sequence
19	1011.6	67.4	5417	6	AX538619 Sequence
20	1006.2	67.0	4402	6	E30219 Shortened d
21	922.2	61.4	13575	5	GCDYS
22	807	53.8	4402	6	E30220
23	715.6	47.7	2654	5	FSCDYSTRO
24	671.4	44.7	4075	6	E30221
25	667.2	44.5	3747	6	E30218 Shortened d
26	613.6	40.9	3163	6	E30223
27	598.2	39.9	3521	5	AF339031
28	555.6	37.0	10705	10	RNAJ2967
29	554.2	36.9	11096	6	AX538584
30	554.2	36.9	11096	10	MMV12229
31	554	36.9	3161	10	MMGUTRPH
32	547.4	36.5	6045	6	A63605
33	547.4	36.5	6045	6	AR281528 Sequence
34	547.4	36.5	6053	6	AX107972 Sequence
35	547.4	36.5	10302	6	AX538583
36	547.4	36.5	10302	9	HSMUFS
37	547.4	36.5	10320	6	A63607
38	547.4	36.5	10320	6	AR281529 Sequence
39	528.4	35.2	3073	10	AF195788
40	528.4	35.2	3140	10	AF195787
41	527.4	35.1	3499	9	HSU43519
42	527.4	35.1	5106	6	AX552248
43	505	33.6	2005	6	AX817318
44	505	33.6	2110	9	HUMDMDX
45	505	33.6	4658	9	BC028720 Homo sapi

ALIGNMENTS

RESULT 1
AX538620
LOCUS AX538620
DEFINITION Sequence 40 from Patent WO0229056.
ACCESSION AX538620
VERSION AX538620.1 GI:25271166
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Parent: WO 0229056-A 40 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

linear PAT 23-NOV-2002

```
FEATURES
  source
    Location/Qualifiers
      1..5339
        /organism="synthetic construct"
        /mol_type="unassigned DNA"
        /db_xref="taxon:32630"
        /note="Synthetic"

ORIGIN
  Query Match      88.5%; Score 1328.8; DB 6; Length 5339;
  Best Local Similarity 99.5%; Pred. No. 0;
  Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 162 GGGTCTTTTACAGACAGTTCTGACAGTGGAGCGTCTGCACCTTTCTCTCAGGAAC 221
DB 1875 GGTCCCATTTTGAAGCCAGTCTGACCAAGTGGAGCGTCTGCACCTTTCTCTCAGGAAC 1934
QY 222 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGCG 281
DB 1935 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGCG 1994
QY 282 ACTTTCAGAGTTTCAAGACGAAACGATGATACATAGGGCCCTTCAAGAGGGAATTCAGAA 341
DB 1995 ACTTTCAGAGTTTCAAGACGAAACGATGATACATAGGGCCCTTCAAGAGGGAATTCAGAA 2054
QY 342 CTAAGAACCTGTAACTCATGACTACTCTTGAGACTGTACGAATATTCTTGACAGCAGC 401
DB 2055 CTAAGAACCTGTAACTCATGACTACTCTTGAGACTGTACGAATATTCTTGACAGCAGC 2114
QY 402 CTTTGAAGGACTAGAGAACTCTTACCAGAGCCGAGAGAGTGCCTCTCTGAGGAGAGAG 461
DB 2115 CTTTGAAGGACTAGAGAACTCTTACCAGAGCCGAGAGAGTGCCTCTCTGAGGAGAGAG 2174
QY 462 CCAGAAATGTCACTCGGCTCTTACGAAAGCAGCTGAGAGGTCAATATCTGATGGGAA 521
DB 2175 CCAGAAATGTCACTCGGCTCTTACGAAAGCAGCTGAGAGGTCAATATCTGATGGGAA 2234
QY 522 AATTGAACCTGCACCTCGCTGACTGCGAGAGAAATAGATGAGACCTTGAAGACTCC 581
DB 2235 AATTGAACCTGCACCTCGCTGACTGCGAGAGAAATAGATGAGACCTTGAAGACTCC 2294
QY 582 AGAACTTCAAGAGGCCAGGATGAGCTGAGCTCAAGCTGCGCCAAAGCTGAGGTGATCA 641
DB 2295 AGAACTTCAAGAGGCCAGGATGAGCTGAGCTCAAGCTGCGCCAAAGCTGAGGTGATCA 2354
QY 642 AGGATCTTGGACCGCGTGGCGATCTCTCTCAATTCATCTCTCCAGATCACTTCGAGA 701
DB 2355 AGGATCTTGGACCGCGTGGCGATCTCTCTCAATTCATCTCTCCAGATCACTTCGAGA 2414
QY 702 AAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGCCACGTCATG 761
DB 2415 AAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGCCACGTCATG 2474
QY 762 ACCTTGTCTGCGAGCTTACACCTTTGGGCAATTCAGCTCTCAACGATACCTTCAGCACTC 821
DB 2475 ACCTTGTCTGCGAGCTTACACCTTTGGGCAATTCAGCTCTCAACGATACCTTCAGCACTC 2534
QY 822 TGAAGACCTGTAACACAGATGAAAGCTTCTGAGGTGCGCTGCGAGGACCGAGTCAGGC 881
DB 2535 TGAAGACCTGTAACACAGATGAAAGCTTCTGAGGTGCGCTGCGAGGACCGAGTCAGGC 2594
QY 882 AGCTGATGAAGCCCAACAGGACTTTGGTCCAGCATCTCAGACCTTTCTTTCCAGCTCTG 941
DB 2595 AGCTGATGAAGCCCAACAGGACTTTGGTCCAGCATCTCAGACCTTTCTTTCCAGCTCTG 2654
QY 942 TCCAGGGTCCCTGGGAGAGAGCATCTCGCCAAACAAAGTGCCTTACTATATCAACACG 1001
DB 2655 TCCAGGGTCCCTGGGAGAGAGCATCTCGCCAAACAAAGTGCCTTACTATATCAACACG 2714
QY 1002 AGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGCTTTTAGCTG 1061
DB 2715 AGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGCTTTTAGCTG 2774
QY 1062 ACCTGAATATGTCAGATTCTCAGCTTATAGACTGCCATGAACTCCGAAGACTGCAGA 1121
DB 2775 ACCTGAATATGTCAGATTCTCAGCTTATAGACTGCCATGAACTCCGAAGACTGCAGA 2834
QY 1122 AGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCTGATGTGATGATGAGCAGACA 1181
DB 2835 AGGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCTGCTGATGTGATGATGAGCAGACA 2894
QY 1182 ACCTCAAGCAAAATGACCAAGCCCATGGATATCTCTGAGATTAATTAATTTGATGACCACTA 1241
DB 2895 ACCTCAAGCAAAATGACCAAGCCCATGGATATCTCTGAGATTAATTAATTTGATGACCACTA 2954
QY 1242 TTTATGACCGCTGGAGCAAGACAGACCAATTTGGTCAACCTCCCTCTCTGCGTGGATA 1301
DB 2955 TTTATGACCGCTGGAGCAAGACAGACCAATTTGGTCAACCTCCCTCTCTGCGTGGATA 3014
QY 1302 TGTGTCTGAATCGCTGCTGCTGAATGTTTATGATACGGACGAAACAGGAGGATCCGTGTCC 1361
DB 3015 TGTGTCTGAATCGCTGCTGCTGAATGTTTATGATACGGACGAAACAGGAGGATCCGTGTCC 3074
QY 1362 TGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGAT 1421
DB 3075 TGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGAT 3134
QY 1422 ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTACAGCGCAGGCTGGCGCTCC 1481
DB 3135 ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTACAGCGCAGGCTGGCGCTCC 3194
QY 1482 TTCTGCAATGTTCTATCCAA 1501
DB 3195 TTCTGCAATGTTCTATCCAA 3214

RESULT 2
AX538621
LOCUS AX538621 5462 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 41 from Patent WO0229056.
ACCESSION AX538621
VERSION AX538621.1 GI:25271168
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 Chamberlain, J.S. and Harper, S.Q.
AUTHORS Mini-dystrophin nucleic acid and peptide sequences
TITLE Patent: WO 0229056-A 41 11-APR-2002;
JOURNAL THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
  source
    Location/Qualifiers
      1..5462
        /organism="synthetic construct"
        /mol_type="unassigned DNA"
        /db_xref="taxon:32630"
        /note="Synthetic"

ORIGIN
  Query Match      88.5%; Score 1328.8; DB 6; Length 5462;
  Best Local Similarity 99.5%; Pred. No. 0;
  Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 162 GGGTCTTTTACAGACAGTTCTGACAGTGGAGCGTCTGCACCTTTCTCTCAGGAAC 221
DB 1998 GGTCCCATTTTGAAGCCAGTCTGACCAAGTGGAGCGTCTGCACCTTTCTCTCAGGAAC 2057
QY 222 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGCG 281
DB 2058 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGCG 2117
QY 282 ACTTTCAGAGTTTCAAGACGAAACGATGATACATAGGGCCCTTCAAGAGGGAATTCAGAA 341
DB 2118 ACTTTCAGAGTTTCAAGACGAAACGATGATACATAGGGCCCTTCAAGAGGGAATTCAGAA 2177
QY 342 CTAAGAACCTGTAACTCATGACTACTCTTGAGACTGTACGAATATTCTTGACAGCAGC 401
```

Db	2178	CTAAGAACCCTGTAATCATGAGTACTCTTGAGACTGTACGAATAATTTCTGACAGAGCAGC	2237
QY	402	CTTTGGAAGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCTCTGAGGAGAG	461
Db	2238	CTTTGGAAGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCTCTGAGGAGAG	2297
QY	462	CCAGAAATCTCACTCGGCTTTCTACGAAGACAGGCTGAGAGGTCAATCTAGTGGGAAA	521
Db	2298	CCAGAAATCTCACTCGGCTTTCTACGAAGACAGGCTGAGAGGTCAATCTAGTGGGAAA	2357
QY	522	AATTGAACCTGCACTCCGCTGAGTGGCAGAGAAAATAGATAGACCCCTTGAAGACTCC	581
Db	2358	AATTGAACCTGCACTCCGCTGAGTGGCAGAGAAAATAGATAGACCCCTTGAAGACTCC	2417
QY	582	AGGAACCTTCAAGAGGCCAGGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGTGAATCA	641
Db	2418	AGGAACCTTCAAGAGGCCAGGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGTGAATCA	2477
QY	642	AGGGATCTTGGCAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCGAGA	701
Db	2478	AGGGATCTTGGCAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCGAGA	2537
QY	702	AGTCAAGGCACTTCAGAGGAAATTCGCTCTGAAAGAGAACGTGAGCCACGTCAATG	761
Db	2538	AGTCAAGGCACTTCAGAGGAAATTCGCTCTGAAAGAGAACGTGAGCCACGTCAATG	2597
QY	762	ACCTTGCTGCCAGCTTACCACTTTTGGGCACTCAGCTCTCACGGTATAACCTCAGACTC	821
Db	2598	ACCTTGCTGCCAGCTTACCACTTTTGGGCACTCAGCTCTCACGGTATAACCTCAGACTC	2657
QY	822	TGGAAGACCTTGAAACAGACAGTGAAGCTTCTGAGGTGCGCTCGAGGACCGAGTCAGC	881
Db	2658	TGGAAGACCTTGAAACAGACAGTGAAGCTTCTGAGGTGCGCTCGAGGACCGAGTCAGC	2717
QY	882	AGTGTGATGAAGCCACAGGCACTTTGGTCCAGCATCTCAGCACTTTTCTTCCAGCTCG	941
Db	2718	AGTGTGATGAAGCCACAGGCACTTTGGTCCAGCATCTCAGCACTTTTCTTCCAGCTCG	2777
QY	942	TCCAGGCTCCCTGGGAGAGACCATCTCGCCAAACAAAGTGCCCTACTATATCAACCAG	1001
Db	2778	TCCAGGCTCCCTGGGAGAGACCATCTCGCCAAACAAAGTGCCCTACTATATCAACCAG	2837
QY	1002	AGACTCAAACTTGTCTGGGAGCCATCCCAAAATGACAGAGCTTACCAGTCTTTAGCTG	1061
Db	2838	AGACTCAAACTTGTCTGGGAGCCATCCCAAAATGACAGAGCTTACCAGTCTTTAGCTG	2897
QY	1062	ACCTGAATAATGTGAGTCTCAGCTTATAGGACTGCGCAAACTCCGAAAGACTGCAGA	1121
Db	2898	ACCTGAATAATGTGAGTCTCAGCTTATAGGACTGCGCAAACTCCGAAAGACTGCAGA	2957
QY	1122	AGGCCCTTTGCTTGGATCTCTGAGCTGTGAGTGTGATGCTTGGAGCAGACA	1181
Db	2958	AGGCCCTTTGCTTGGATCTCTGAGCTGTGAGTGTGATGCTTGGAGCAGACA	3017
QY	1182	ACCTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTAATTAATTTTGACCACTA	1241
Db	3018	ACCTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTAATTAATTTTGACCACTA	3077
QY	1242	TTTATGACCGCTCGAGCAAGAGACAACTTTGGTCAAGCTCCCTCTCTGCGTGGATA	1301
Db	3078	TTTATGACCGCTCGAGCAAGAGACAACTTTGGTCAAGCTCCCTCTCTGCGTGGATA	3137
QY	1302	TGCTCTGAACTGCTGCTGAACTTTATGATACGGGACGAAACAGGAGGATCCGTGTCC	1361
Db	3138	TGCTCTGAACTGCTGCTGAACTTTATGATACGGGACGAAACAGGAGGATCCGTGTCC	3197
QY	1362	TGCTTTTAAACTGGCATATTTCCCTGTGTAAAGACACATTTTGGAAAGACAAGTACAGAT	1421
Db	3198	TGCTTTTAAACTGGCATATTTCCCTGTGTAAAGACACATTTTGGAAAGACAAGTACAGAT	3257
QY	1422	ACCTTTTCAAGCAAGTGGCAATTTTCAAGCAATTTTGGACCGGAGGCTGGCCCTCC	1481
Db	3258	ACCTTTTCAAGCAAGTGGCAATTTTCAAGCAATTTTGGACCGGAGGCTGGCCCTCC	3317

QY

1482

TTCTGCATGATTTCTATCCAA

1501

Db

3318

TTCTGCATGATTTCTATCCAA

3337

RESULT 3

AR304538

LOCUS

Sequence 1 from patent US 6544786.

DEFINITION

AR304538

ACCESSION

AR304538.1

VERSION

GI:31693691

KEYWORDS

Unknown.

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 5952)

AUTHORS

Xiao.X. and Liu.P.X.

TITLE

Method and vector for producing and transferring trans-spliced peptides

JOURNAL

Patent: US 6544786-A 1 08-APR-2003;

FEATURES

Location/Qualifiers

1..5952

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match

88.5%; Score 1328.8; DB 6; Length 5952;

Best Local Similarity

99.5%; Pred. No. 0;

Matches 1333; Conservative

0; Mismatches

7; Indels

0; Gaps

0;

QY

162

GGGTCTTTTACAAGACAGATTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCTCTGAGGAAAC

221

Db

3284

GGTCCATTTGGAGCCAGTTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCTCTGAGGAAAC

3343

QY

222

TTCTGCTGCTCAGCTGAAAGATGATGAATTAAGCCGCGAGCGACCTATTGGAGCG

281

Db

3344

TTCTGCTGCTCAGCTGAAAGATGATGAATTAAGCCGCGAGCGACCTATTGGAGCG

3403

QY

282

ACTTTCCAGCAGTTTCAGAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAAA

341

Db

3404

ACTTTCCAGCAGTTTCAGAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAAA

3463

QY

342

CTAAAGAACCTGTATCATGTAGTACTCTTGAGACTGTACGAATATTTTGTGACAGCAGC

401

Db

3464

CTAAAGAACCTGTATCATGTAGTACTCTTGAGACTGTACGAATATTTTGTGACAGCAGC

3523

QY

402

CTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCGAGAGCTGCTCTCTGAGGAGAG

461

Db

3524

CTTTGGAAGGACTAGAGAAACTCTACAGGAGCCCGAGAGCTGCTCTCTGAGGAGAG

3583

QY

462

CCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAA

521

Db

3584

CCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAA

3643

QY

522

AATTGAACCTGCACTCCGCTGAGTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC

581

Db

3644

AATTGAACCTGCACTCCGCTGAGTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC

3703

QY

582

AGGAACCTTCAAGAGCCAGGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGTGAATCA

641

Db

3704

AGGAACCTTCAAGAGCCAGGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGTGAATCA

3763

QY

642

AGGGATCTCTGGCAGCCGCTGGGCGATCTCCTCATTGACTCTCTCTCAAGATCACTCGAGA

701

Db

3764

AGGGATCTCTGGCAGCCGCTGGGCGATCTCCTCATTGACTCTCTCTCAAGATCACTCGAGA

3823

QY

702

AGTCAAGGCACTTCAGAGGAAAATTCGCTCTGAAAGAGAACGCTGAGCCACGTCAATG

761

Db

3824

AGTCAAGGCACTTCAGAGGAAAATTCGCTCTGAAAGAGAACGCTGAGCCACGTCAATG

3883

QY

762

ACCTTCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTTCCCGTATTAACCTCAGCACTC

821

```

Db 3884 ACCTTGCTGCCAGCTTACCACTTTGGGCACTTACGCTCTCACCGTATATAAAGCTCAGCACTC 3943
Qy 822 TGGAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTACGCG 881
Db 3944 TGGAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTACGCG 4003
Qy 882 AGCTGCATGAAGCCACACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTG 941
Db 4004 AGCTGCATGAAGCCACACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTG 4063
Qy 942 TCCAGGTCCTCGGAGAGCCATCTCGCAAAACAAAGTCCCTACTATATCAACCAAG 1001
Db 4064 TCCAGGTCCTCGGAGAGCCATCTCGCAAAACAAAGTCCCTACTATATCAACCAAG 4123
Qy 1002 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGCTG 1061
Db 4124 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGCTG 4183
Qy 1062 ACCTGAATAATGTCCAGATTTCTCAGCTTATAGGACTGCCATGAATACTCCGAAGACTGCAGA 1121
Db 4184 ACCTGAATAATGTCCAGATTTCTCAGCTTATAGGACTGCCATGAATACTCCGAAGACTGCAGA 4243
Qy 1122 AGGCCCCCTTGGTGGTCTCTTGGAGCTCTCAGTGCATGTGATGCTTGGAGCAGACAGA 1181
Db 4244 AGGCCCCCTTGGTGGTCTCTTGGAGCTCTCAGTGCATGTGATGCTTGGAGCAGACAGA 4303
Qy 1182 ACCTCAAGCAAAATGACAGCCATGATATCTCGAGATTAATTTGTTGACCACTA 1241
Db 4304 ACCTCAAGCAAAATGACAGCCATGATATCTCGAGATTAATTTGTTGACCACTA 4363
Qy 1242 TTATGACCGCTGGAGCAGACAGCAAAATTTGTTCAAGTCCCTCTCTCGCTGGATA 1301
Db 4364 TTATGACCGCTGGAGCAGACAGCAAAATTTGTTCAAGTCCCTCTCTCGCTGGATA 4423
Qy 1302 TGTGCTGAAGTGGCTCTCAATGTTTATGATAGGAGCGAAGCAGGAGGATCCGTTGCC 1361
Db 4424 TGTGCTGAAGTGGCTCTCAATGTTTATGATAGGAGCGAAGCAGGAGGATCCGTTGCC 4483
Qy 1362 TGCTTTTAAATCGGCATCATTTCCCTGTGTAAGCAGCATTTTGAAGCAAAAGTACAGAT 1421
Db 4484 TGCTTTTAAATCGGCATCATTTCCCTGTGTAAGCAGCATTTTGAAGCAAAAGTACAGAT 4543
Qy 1422 ACCTTTTCAAGCAAGTGGCAGTTCACAGGATTTTGTACAGCAGCAGGCTGGGCTCC 1481
Db 4544 ACCTTTTCAAGCAAGTGGCAGTTCACAGGATTTTGTACAGCAGCAGGCTGGGCTCC 4603
Qy 1482 TTCTGCATGATTTATCCAA 1501
Db 4604 TTCTGCATGATTTATCCAA 4623

```

```

RESULT 4
AX114289
LOCUS
DEFINITION
Sequence 1 from Patent WO0129243.
AX114289
ACCESSION
AX114289.1 GI:14031259
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Paul X.L. and Xiao X.
TITLE
Method and vector for producing and transferring trans -spliced
Peptides
JOURNAL
Patent: WO 0129243-A 1 26-APR-2001;
DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH
SYSTEM OF HIGHER EDUCATION (US)
FEATURES
Location/Qualifiers
1..5952
/organism="Homo sapiens"
/mol_type="unassigned DNA"
source

```

```

/db_xref="taxon:9606"
misc_feature
2897..2898
/Note="S4 junction site"
misc_feature
3198..3199
/Note="S2 junction site"
ORIGIN
Query Match 88.5%; Score 1328.8; DB 6; Length 5952;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 162 GGGTTCTTTTCAAGACAGATTTCTGACAGTGGAGCGCTCTGACCTTTCTCTGAGGAGC 221
Db 3284 GGTCCCAATTTGAGAGCCAGTTCTGACCAAGTGAAGCGTCTGACCTTTCTCTGAGGAGC 3343
Qy 222 TTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACTTATTGAGGCG 281
Db 3344 TTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACTTATTGAGGCG 3403
Qy 282 ACTTTCCAGCAGTTTCAGNAGCAGAACGATGTACATAGGGCTTTCAGAGGGAATTTGAAA 341
Db 3404 ACTTTCCAGCAGTTTCAGNAGCAGAACGATGTACATAGGGCTTTCAGAGGGAATTTGAAA 3463
Qy 342 CTAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 401
Db 3464 CTAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 3523
Qy 402 CTTTGGAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCCTCTGAGGAGAGAG 461
Db 3524 CTTTGGAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCCTCTGAGGAGAGAG 3583
Qy 462 CCCAGAAATGTCACCTCGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTGAGTGGGAAA 521
Db 3584 CCCAGAAATGTCACCTCGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTGAGTGGGAAA 3643
Qy 522 AATTGAACCTTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAGACTCC 581
Db 3644 AATTGAACCTTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAGACTCC 3703
Qy 582 AGGAACCTTCAAGAGGCCACAGATGAGCTGACCTCAGCTGCGCCAGCTGAGGTGATCA 641
Db 3704 AGGAACCTTCAAGAGGCCACAGATGAGCTGACCTCAGCTGCGCCAGCTGAGGTGATCA 3763
Qy 642 AGGATCTCTGCGACCCGCTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGA 701
Db 3764 AGGATCTCTGCGACCCGCTGGCGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGA 3823
Qy 702 AAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTCGAGCAGTCAATG 761
Db 3824 AAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTCGAGCAGTCAATG 3883
Qy 762 ACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTGAGCTCTCACCGTATAAAGCTCAGCACTC 821
Db 3884 ACCTTGCTCGCCAGCTTACCACCTTTGGGCATTTGAGCTCTCACCGTATAAAGCTCAGCACTC 3943
Qy 822 TGGAGACCTGAAACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTACGCG 881
Db 3944 TGGAGACCTGAAACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTACGCG 4003
Qy 882 AGCTGCATGAAGCCACACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTG 941
Db 4004 AGCTGCATGAAGCCACACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTG 4063
Qy 942 TCCAGGTCCTCGGAGAGCCATCTCGCAAAACAAAGTCCCTACTATATCAACCAAG 1001
Db 4064 TCCAGGTCCTCGGAGAGCCATCTCGCAAAACAAAGTCCCTACTATATCAACCAAG 4123
Qy 1002 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGGCTCTACCACTCTTTAGCTG 1061
Db 4124 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGGCTCTACCACTCTTTAGCTG 4183
Qy 1062 ACCTGAATAATGTCCAGATTTCTCAGCTTATAGGACTGCCATGAATACTCCGAAGACTGCAGA 1121

```

Db 4184 ACCTGATTAATGTCAGATTCTCAGCTTATAGACTGCCATGAACTCCGAGACTGCAGA 4243
QY 1122 AGGCCCTTTGCTTGGATCTCTTGGAGCTCTCAGCTGCATGTCATGTCCTTGGACCAAGACA 1181
Db 4244 AGGCCCTTTGCTTGGATCTCTTGGAGCTCTCAGCTGCATGTCATGTCCTTGGACCAAGACA 4303
QY 1182 ACCTCAAGCAAAATGACCAAGCCCATGATATCTCGAGATTAATTAATGTTTGGACCACTA 1241
Db 4304 ACCTCAAGCAAAATGACCAAGCCCATGATATCTCGAGATTAATTAATGTTTGGACCACTA 4363
QY 1242 TTATGACCGCTGGAGCAAGACCAACAATTTGGTCAACGTCCTCTCTCGCTGGATA 1301
Db 4364 TTATGACCGCTGGAGCAAGACCAACAATTTGGTCAACGTCCTCTCTCGCTGGATA 4423
QY 1302 TGTGCTGCACTGCTGCTGCTGAATGTTTATACCGGACGACACAGGAGGATCCGTGTC 1361
Db 4424 TGTGCTGCACTGCTGCTGCTGAATGTTTATACCGGACGACACAGGAGGATCCGTGTC 4483
QY 1362 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGCAAGTACAGAT 1421
Db 4484 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGCAAGTACAGAT 4543
QY 1422 ACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACAGCGCAGCTGGGCTCC 1481
Db 4544 ACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACAGCGCAGCTGGGCTCC 4603
QY 1482 TTCTGCATGATCTATCCCA 1501
Db 4604 TTCTGCATGATCTATCCCA 4623

RESULT 5
AX538622
LOCUS AX538622 8689 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 42 from Patent WO0229056.
ACCESSION AX538622
VERSION AX538622.1 GI:25271171
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 42 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source
1..8689
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"

ORIGIN
Query Match 88.5%; Score 1328.8; DB 6; Length 8689;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 162 GGGTCTCTTTTACAGACAGTCTGACCAAGTGGAGCGCTGTCACCTTTCTCGAGGAAC 221
Db 3330 GGTCCCAATTTGGAAGCCAGTCTGACCAAGTGGAGCGCTGTCACCTTTCTCGAGGAAC 3389
QY 222 TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCG 281
Db 3390 TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCG 3449
QY 282 ACTTTTCAGAGATTTCAGAGCAGAACCATGTACATAGGCGCTTCAAGAGGGAATTGAAA 341
Db 3450 ACTTTTCAGAGATTTCAGAGCAGAACCATGTACATAGGCGCTTCAAGAGGGAATTGAAA 3509
QY 342 CTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATTTTCTGACAGAGCAGC 401
Db 3510 CTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATTTTCTGACAGAGCAGC 3569

QY 402 CTTTGGAGGACTAGAGAACTCTTACCAGAGCCCGAGAGAGCTGCCTCTCTGAGAGAGAG 461
Db 3570 CTTTGGAGGACTAGAGAACTCTTACCAGAGCCCGAGAGAGCTGCCTCTCTGAGAGAGAG 3629
QY 462 CCCAGAATGTCACTCGGCTTCTACGAAAGCAGCGTGAAGAGTCAATATCTAGTGGGAAA 521
Db 3630 CCCAGAATGTCACTCGGCTTCTACGAAAGCAGCGTGAAGAGTCAATATCTAGTGGGAAA 3689
QY 522 AATTGAACCTGCACTCGGCTGATCGGAGAGAAATAGATGAGACCTTTGAAAGACTCC 581
Db 3690 AATTGAACCTGCACTCGGCTGATCGGAGAGAAATAGATGAGACCTTTGAAAGACTCC 3749
QY 582 AGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGTGTATCA 641
Db 3750 AGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGTGTATCA 3809
QY 642 AGGATCTCTGCAACCGCTGGCGGATCTCTCATTTGACTCTCTCAAGATCACTCCGAGA 701
Db 3810 AGGATCTCTGCAACCGCTGGCGGATCTCTCATTTGACTCTCTCAAGATCACTCCGAGA 3869
QY 702 AAGTCAAGGCACTTTCGAGGAGAAATTCGCTCTGAAAGAGAACTGAGCCACCTCAATG 761
Db 3870 AAGTCAAGGCACTTTCGAGGAGAAATTCGCTCTGAAAGAGAACTGAGCCACCTCAATG 3929
QY 762 ACCTTGTCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAACCCTAGACTC 821
Db 3930 ACCTTGTCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAACCCTAGACTC 3989
QY 822 TGGAGACCTGAACACACAGATGGAAGCTTCTGAGGTTGGCGCTGAGACCGAGTCAAGGC 881
Db 3990 TGGAGACCTGAACACACAGATGGAAGCTTCTGAGGTTGGCGCTGAGACCGAGTCAAGGC 4049
QY 882 AGCTGCATGAAGCCACAGGAGCTTTGGTCCAGATCTCAGACATTTCTTTCCACGTGTG 941
Db 4050 AGCTGCATGAAGCCACAGGAGCTTTGGTCCAGATCTCAGACATTTCTTTCCACGTGTG 4109
QY 942 TCCAGGTCCTTGGAGAGAGCCATCTCGCCAAACAAAGTCCCTCTATATCAACACG 1001
Db 4110 TCCAGGTCCTTGGAGAGAGCCATCTCGCCAAACAAAGTCCCTCTATATCAACACG 4169
QY 1002 AGACTCAAAACAACTTGTGGGACCATCCAAATGACAGAGCTTACCAGTCTTTAGCTG 1061
Db 4170 AGACTCAAAACAACTTGTGGGACCATCCAAATGACAGAGCTTACCAGTCTTTAGCTG 4229
QY 1062 ACCTGAATTAATGTCAAGTCTCAGCTTATAGACTGCATGAACTCGAAGACTGCAGA 1121
Db 4230 ACCTGAATTAATGTCAAGTCTCAGCTTATAGGACTGCCATGAACTCCGAAAGACTGCAGA 4289
QY 1122 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCTTGGACCAAGACA 1181
Db 4290 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCTTGGACCAAGACA 4349
QY 1182 ACCTCAAGCAAAATGACCAAGCCCATGGATATCTCTGAGATTAATTAATTTGTTGACCACTA 1241
Db 4350 ACCTCAAGCAAAATGACCAAGCCCATGGATATCTCTGAGATTAATTAATTTGTTGACCACTA 4409
QY 1242 TTATGACCGCTGAGCAAGACACACAAATTTGGTCAACGCTCTCTCGCTGGATA 1301
Db 4410 TTATGACCGCTGAGCAAGACACAAATTTGGTCAACGCTCTCTCGCTGGATA 4469
QY 1302 TGTGCTGAACCTGGCTGCTGAATGTTTATGATAGGACGAAACAGGAGGATCCGTGTCC 1361
Db 4470 TGTGCTGAACCTGGCTGCTGAATGTTTATGATAGGACGAAACAGGAGGATCCGTGTCC 4529
QY 1362 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGCAAGTACAGAT 1421
Db 4530 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGCAAGTACAGAT 4589
QY 1422 ACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCGCCAGGCTGGGCTCC 1481
Db 4590 ACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCGCCAGGCTGGGCTCC 4649

ORIGIN									
Query Match 88.5%; Score 1328.8; DB 6; Length 12057;									
Best Local Similarity 99.5%; Pred. No. 0;									
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;									
Qy	162	GGGTTCTTTTACAGACAGTCTGACACAGTGGAGCGTCTGACACCTTTCTCTGCAGGAAC	221						
Db	8598	GGTCCCATTTGGAAGCCAGTCTGACACAGTGGAGCGTCTGACACCTTTCTCTGCAGGAAC	8657						
Qy	222	TTCTGTTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGG	281						
Db	8658	TTCTGTTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGG	8717						
Qy	282	ACTTTCAGAGTTCAGAGCAGAAAGATGTAATAGGCGCTTCAAGAGGGAATCAAAA	341						
Db	8718	ACTTTCAGAGTTCAGAGCAGAAAGATGTAATAGGCGCTTCAAGAGGGAATCAAAA	8777						
Qy	342	CTAAGAACCTGTAATCATGAGTACTCTTCAGACTGTACGAATATTCTGACAGACAGC	401						
Db	8778	CTAAGAACCTGTAATCATGAGTACTCTTCAGACTGTACGAATATTCTGACAGACAGC	8837						
Qy	402	CTTTGGAAGGACTAGAGAACTCTTACAGAGAGCCAGAGAGCTGCCTCTGAGGAGAG	461						
Db	8838	CTTTGGAAGGACTAGAGAACTCTTACAGAGAGCCAGAGAGCTGCCTCTGAGGAGAG	8897						
Qy	462	CCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTGAGTGGGAA	521						
Db	8898	CCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTGAGTGGGAA	8957						
Qy	522	AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCC	581						
Db	8958	AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCC	9017						
Qy	582	AGGAATTCAGAGGACCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	641						
Db	9018	AGGAATTCAGAGGACCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	9077						
Qy	642	AGGATCTCTGGCAGCCGCTGGGCGATCTCCTCATTTGACTCTCTCCAGATCACTCGAGA	701						
Db	9078	AGGATCTCTGGCAGCCGCTGGGCGATCTCCTCATTTGACTCTCTCCAGATCACTCGAGA	9137						
Qy	702	AAGTCAAGGCACTTCGAGGAGAAATTCGGCTCTGAAGAGAAAGTGAAGCAGCTCAATG	761						
Db	9138	AAGTCAAGGCACTTCGAGGAGAAATTCGGCTCTGAAGAGAAAGTGAAGCAGCTCAATG	9197						
Qy	762	ACCTTGCTGCCAGCTTACCACTTTGGGCACTTCAAGTCTCAGCGTATAAAGTCAAGCTC	821						
Db	9198	ACCTTGCTGCCAGCTTACCACTTTGGGCACTTCAAGTCTCAGCGTATAAAGTCAAGCTC	9257						
Qy	822	TGGAAGACCTGAACACAGATGGAAGCTTCTGCAGGTGGCGCTCGAGACCCGAGTCAGGC	881						
Db	9258	TGGAAGACCTGAACACAGATGGAAGCTTCTGCAGGTGGCGCTCGAGACCCGAGTCAGGC	9317						
Qy	882	AGCTGCATGAAGCCACAGGACCTTGGTCCAGCATCTCAGCATCTTCTTCCACGCTCG	941						
Db	9318	AGCTGCATGAAGCCACAGGACCTTGGTCCAGCATCTCAGCATCTTCTTCCACGCTCG	9377						
Qy	942	TCCAGGCTCCCTGGGAGAGACCATCTCGCCAAAGTGCCTACTATATCAACCCAGC	1001						
Db	9378	TCCAGGCTCCCTGGGAGAGACCATCTCGCCAAAGTGCCTACTATATCAACCCAGC	9437						
Qy	1002	AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTAGCTG	1061						
Db	9438	AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTAGCTG	9497						
Qy	1062	ACCTGAATATGTACAGATCTCAGCTTATAGGACTGCCATGAAGTCCGAGACTCGAGA	1121						
Db	9498	ACCTGAATATGTACAGATCTCAGCTTATAGGACTGCCATGAAGTCCGAGACTCGAGA	9557						
Qy	1122	AGGCCCTTTGTTGATCTCTTGAAGCTGTGATGATGATGATGATGATGATGATGATGATG	1181						
Db	9558	AGGCCCTTTGTTGATCTCTTGAAGCTGTGATGATGATGATGATGATGATGATGATGATG	9617						

QY	1182	ACCTCAAGCAAAATGACAGCCCATGATATCTCTGCAGATTATTAATTTGTTTACCACTA	1244						
Db	9618	ACCTCAAGCAAAATGACAGCCCATGATATCTCTGCAGATTATTAATTTGTTTACCACTA	9677						
QY	1242	TTTATGACCGCTGGAGCAAGACAGACCAAAATTTGTTCAACGTCCTCTCTCGTGTGGATA	1301						
Db	9678	TTTATGACCGCTGGAGCAAGACAGACCAAAATTTGTTCAACGTCCTCTCTCGTGTGGATA	9737						
QY	1302	TGTGTTCTAACTGCTGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGAGATCCGTGTCC	1361						
Db	9738	TGTGTTCTAACTGCTGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGAGATCCGTGTCC	9797						
QY	1362	TGTCCTTTTAAACTGGGATCATTTTCCCTGTTAAAGCACATTTTGGAAAGCAAACTACAGAT	1421						
Db	9798	TGTCCTTTTAAACTGGGATCATTTTCCCTGTTAAAGCACATTTTGGAAAGCAAACTACAGAT	9857						
QY	1422	ACCTTTTCAAGCAAGTGCAAGTTCAACAGATTTTGTGACGAGCGAGGCTGGGCTCTCC	1481						
Db	9858	ACCTTTTCAAGCAAGTGCAAGTTCAACAGATTTTGTGACGAGCGAGGCTGGGCTCTCC	9917						
QY	1482	TTCTGCATGATTCTATCCAA	1501						
Db	9918	TTCTGCATGATTCTATCCAA	9937						
RESULT 8									
HSDMDR									
LOCUS	Human mRNA for dystrophin.			12446 bp	mRNA	linear	PRI 12-SEP-1993		
DEFINITION	X14298								
ACCESSION	X14298.1 GI:30845								
VERSION	Dmd gene; Duchenne muscular dystrophy; dystrophin.								
KEYWORDS	Homo sapiens (human)								
SOURCE	Homo sapiens								
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
REFERENCE	1 (bases 1 to 12446)								
AUTHORS	Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and Davies,K.E.								
TITLE	Two human cDNA molecules coding for the Duchenne muscular dystrophy (DMD) locus are highly homologous								
JOURNAL	Nucleic Acids Res. 17 (13), 5391 (1989)								
MEDLINE	89345106								
PUBMED	2668885								
REFERENCE	2 (bases 1 to 12446)								
AUTHORS	Rosenthal,A.								
TITLE	Direct Submission								
JOURNAL	Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften der DDR, Zentralinstitut fuer Molekularbiologie, Robert-Rössle Str.10, 1115 Berlin Buch, DDR								
COMMENT	see also M18533 and M20250 for Dmd seqs.; discrepancies compared to M18533 cDNA were located at X14298 pos. 496, 1772, 1965, 2449, 3687, 4229, 4504, 5075, 5332, 5630 and 7194.								
FEATURES	Location/Qualifiers								
source	1..12446								
	/organism="Homo sapiens"								
	/mol_type="mRNA"								
	/db_xref="taxon:9606"								
	/chromosome="X chromosomal, Xp21."								
	/tissue type="muscle"								
	/dev stage="fetal and adult."								
	99..11156								
	/note="unnamed protein product; dystrophin (AA 1 - 3685) "								
	/codon_start=1								
	/protein_id="CAA32479.1"								
	/db_xref="GI:30846"								
	/db_xref="GOA:P11532"								
	/db_xref="SWISS-PROT:P11532"								
	/translation="VLWEEVEDCVBERDYVKQKFTTKVNAQFSKFGKHIEILFSLDQDCRLRLDLGLTCQKLPRKSGSTRVHALNNVKNLRLVONNVLDLVNIGSDIVDGNKHUTLGLIWIILHMQVKNKIMAGLQPTNSEKILLSWRCSTREYFQVNVINFTTSWDLGLNALILHSHRPDLFDWNSVVCQSATORLEKILLVQGLIEKLLDLPDPT								

RESULT 8

HSDMDR Human mRNA for dystrophin. 12446 bp mRNA linear PRI 12-SEP-1993
DEFINITION Human mRNA for dystrophin.
LOCUS X14298
ACCESSION X14298.1 GI:30845
VERSION Dmd gene; Duchenne muscular dystrophy; dystrophin.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 12446)
AUTHORS Rosenthal A., Speer, A., Billwitz, H., Cross, G.S., Forrest, S.M. and Davies, K.E.
TITLE Two human cDNA molecules coding for the Duchenne muscular dystrophy (DMD) locus are highly homologous
JOURNAL Nucleic Acids Res. 17 (13), 5391 (1989)
MEDLINE 89345106
PUBMED 2668885
REFERENCE 2 (bases 1 to 12446)
AUTHORS Rosenthal, A.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften der DDR, Zentralinstitut fuer Molekularbiologie, Robert-Rössle Str.10, 1115 Berlin Buch, DDR
COMMENT see also M18533 and M20250 for Dmd seqs.; discrepancies compared to M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449, 3687, 4229, 4504, 5075, 5332, 5630 and 7194.

FEATURES

source
1..12446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X chromosomal, Xp21."
/tissue_type="muscle"
/dev_stage="fetal and adult."
99..11156
/note="unnamed protein product; dystrophin (AA 1 - 3685)"
/codon_start=1
/protein_id="CAA32479.1"
/db_xref="GI:30845"
/db_xref="GOA:P11532"
/db_xref="SWISS-PROT:P11532"
/translation="MLWEEVEDCYEREDVOKTFTKWNQAFSKFGKHENI.FSDL QDRLRLLEGLTCKLEKFKGSTRHALNNKALRVLQNNVDLVNIGSTIVDG NHKLTGLTWNIIHLWQVNMKNMAGLQPTNSEKILLSWTRSTRYPGVINFT TSWSDGLNALNIHSHRDLFDWNSVWCQSQATQRLEHAFNIARYQLGKLLQPEDV

DTTVPDKKILMYITSLFOVLPOOVSTIEALQEVEMLPFRPPKVTKEBHFQHLHOMHVSQ
 QITVSLAQGVTSPPKFRKSYATYAAAYTTSDTPRSPFPSPHLEAEDEKDFSGSL
 MESVNDIROYTALIEVLSWLAEDTLQAGBEISNDVEVVKQDFHTEGVMMDLTAFH
 QRGVNTIQLQGLTIGTGLKSEDETEVOQMNLLNSWECFLVASMWKSQYNLHRYLM
 DLONKLELNDMLTTEERTRKWEPLGDLIEDLKQVQCHVKQLQDLBOQVQVNV
 SLTHWVVVDSGDHATAALEBOLKVLGDRMANICWTEDEWVLYLQDILLKQRLTE
 EQCLFSAWSEKEDAVNKHITHTFKQONEMSSQLAVLKADLERKQSGMKGLKSIK
 QDITLTKNKEVTOCTEAWLNDFAKWDNLVQLEKSTAOISOAVITTFPSITQTTFW
 ELLVTTTRQIQLVHAQEBELPPPPQKQKQITVDSBEIRKRLDVIDITELHSMITRSEA
 VLOSQFAIRKNSGFDLKQVNAIERAEKFRKLQDASQCALVQMWNEGVNA
 DSIQASEQLNKEWIEFCOLLSERLNWLQYNNIIAPYNOLQLEQWITAEANLWKIQ
 TTTFSEPTALKSOLKICQDVNRSLGLOPOIERLKIQSIALKEKGGKPMFLDADFVAF
 PTHFKOVFSDVQAEKELOITFDTLPMRYOETMSAIRTWQSEFKLSIPQLSVTDY
 EIMQRLGELQALQSSLOBOQSGLYLSTTVKEMSKAPSEISRKTKOSFEETEGRWK
 KLSQLVEHQCEQEQMNKURKIQNHITLUKMWMAEVDVFLKEENFALGDSBELLKQOL
 KQCLLVSDICTIQPSLNSVNEGQIKNEAEPEFASRLJETELKELNTQWDRNCQOYV
 ARKEALGGLEKTVSLQKLSSEHWNTOAESEYLERDFEYKTFDLQKAFEMEKAK
 EEAQCKEAKVLLTVSNGSVIAQAPVQALKELETITNTYQMLCTLNGKCKTLE
 NQRIALAGTLDGWMDELNELETENSEWSEBLSHBAVRBOKLLQSIOASQETNS
 EVMACHILLSYLEKANKWLNVEFKLTKTENIPGGAEBISVLDLSLENLMRHSNDP
 NOLRLAQTLDGWMDELNELETENSEWSEBLSHBAVRBOKLLQSIOASQETNS
 LHLQESLTFIDQLAAYIADKVDAAQMPQEAQKIQSDLTSHSEISLEEMKKNQKGEA
 AQRVLSQIDVAQKQLQDVSMKFRUFQKPAFQRLQESKMILDEWQXHLPALETKEVE
 QEVVQSNHCVNLYLSSEVKEVEMVITGRQIVQKQOTENPKELDERVATLKLHY
 NELGAKYTERKQLEKLSRMRKEMNVLTEWLAATMDLTKRGAEGMPSNDLSE
 VANGKATQIEQKQVHLKXITVEGALKTVLQKKTLEVLKSLNSNWNIAVTSRAE
 EWLNLLEYOKHMETPONVDHITKMTIOADTLLDESEKKPOCKEDVLKRLKAELEND
 IRPKVDSRTOQANIMANHGDKRLVPOISLHNHFAAISHRIKTKGASIPLEKE
 QFNSDIOKLELPEAELOQGNLKEEDFNKDMEDNEGTVKELLQKGNLQKQITDER
 KSEBIKIQKQLQTNALDKLSRRKKALEISHOWYQKQADQDLCLDLDIEKKL
 ASPEPDERKIEIDELQKKEELNVRQAEGLSEDAANAVEPTQIOLSKWRE
 IESFAQRLNLTQVREEMTWMDLPLEIISYVPSYLTETVHVSQALLEVQ
 LLNAPDLCADFELFKOESLXNKIDSLQSSGRDITHSKXTAALQSATPVERVKL
 VAGKATQIEQKQVHLKXITVEGALKTVLQKKTLEVLKSLNSNWNIAVTSRAE
 QEALSDOLFQWENKVMYKDRQGRFDRSEKWERFHYDKIFQWLTSEAEQLKTOI
 PENWEHAKYKYLKELQDGIQRTVVRILNATGEBEIIQSSSKTASILQELGSLNL
 RWBEQKLSRQKRESEKQNLISFQRLNNEFVWLEADNIASIPLFPGKEQQLKE
 KLBQVLLVBEELPRLQIGLQKNETGQVLVSAPISEFQDLKELNKLQTNLQWIKVS
 RALFPEKQETIEIKOLQLEKLEDELEQNLHLLWLSPIRNLQIYVNPQEGPFD
 VKETIEAVQAKQPVNIEISKGQHLKKEPATQPVKRLKEDLSSEKVAVNLRLQRA
 KQDPLAPGLTIGASPTQVTVTPVATKETAISLEMPSSLSSEKVAVNLRLQRA
 ELTMDLSLQVILKQSVRMVQDLEDINEMIIOKQATMDLEORRPOLEELITAEONLK
 NKTNSQBPARTIITDRERIONQWDEQVHQLNRRQQLNEMLDSTOWLEAKESQVIL
 QKAKLESWKEGPTVDALQKAITETQKALQKLRQWQNVNDVANDLAKLRLQSD
 DTRKHMITINASWRSIHKRVSEAEALHETRLQFPDLDEFLAWLETTA
 NVLQDTRKELLEDKGVLEKQMDQLOQETEAHTDVYHNLDSQKILRSLEGS
 DVALLRDLNMMKSELKKSINRSHLEASSDQWKLHLSLQBLVLVWLQKXDEL
 SROAPIGCDPAPVQKNDVHRAFKRELKKEPVIMSTLETVRIFLTEQPLEGLEKLYQ
 EPRELPEERQAVNTRLLKQAEVANTWEKLNHLSADQKIDETLEBLOLEQATD
 ELDLKQRAQVILKQVQPGDILLIDSLQHLEKVKALGEIAPLKENVSHVNDLQRL
 TTIGIQSPNLSLEPDLNTRKWLQVAVEDVRQJHEAHRFGPASQHFSTSVQGP
 WERASIPKVFYINNETQTCWDHPKMTLEYQSLADLNVRFSAYRTANMLVRLOKA
 LCILDSLASDALDOLHNLQNDPMDILQINCLTITYDRLEQHNKLNVPLQV
 MCNLNLNVYDTQRTGRIRVLSFKTGIISLKAHLEDKRYFLFKVQASSTGCDORRL
 GLLLHDSIQIOLPRLGVEASVPGSNIIPSVRSQFQFANNKEIPEALFLDWMRLPQSM
 VMLPVLVAAATKHOAKNICKECPIIIGFEYRSLKHPNDVICOSCFESGRVAKH
 KHYHVMVEYCTPTSGEDVRDPAKLVNKRFRKRYPAKHPRMGLPVQTVLSDGMET
 PVTILNFWPDSAPASPOLSHDTHSRHIEHVASLRAEMENSGNSYLSNPSID
 DEHLLOHYCQSLNQSPUSQPSRPAQILISLESEERBELILADLEARNENLOLEY
 DRILQOHEHKGSLPFSPEMPTSPQSPRDAELIAEAKLLRQHGRLEARNQIILEDH
 NKQESQLHRLQLLPQPAEAKVNGTIVSSPSTLSQRSDSOPMLLRVVGSTQSDSM
 GEEDLLSPQDSTGTGLEVMEQLNNSFPSSRGRTPGKPMREDTVM

ORIGIN

Query Match 88.5%; Score 1328.8; DB 9; Length 12446;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy	162	GGGTTCCTTTTACAGACAGTTCGACCAGTGGAGCGTCTGCACCTTTCTCTGCGAGAAC	221
Db	8488	GGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAGCGTCTGCACCTTTCTCTGCGAGAAC	8547
Qy	222	TTCTGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGAGGCG	281
Db	8548	TTCTGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGAGGCG	8607

Qy	282	ACTTTCCAGCAGTTCCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAAATTGAAAA	341
Db	8608	ACTTTCCAGCAGTTCCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAAATTGAAAA	8667
Qy	342	CTAAAGAACCTGTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC	401
Db	8668	CTAAAGAACCTGTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC	8727
Qy	402	CTTTTGGAAAGGACTAGAGAAAATCTTACCAAGAGCCCGAGAGAGCTGCCTCTCTGAGAGAGAG	461
Db	8728	CTTTTGGAAAGGACTAGAGAAAATCTTACCAAGAGCCCGAGAGAGCTGCCTCTCTGAGAGAGAG	8787
Qy	462	CCAGAAATGTCACTCGGCTTCTAGAAAGAGAGGCTGAGGAGGTCAATACTGAGTGGGAAA	521
Db	8788	CCAGAAATGTCACTCGGCTTCTAGAAAGAGAGGCTGAGGAGGTCAATACTGAGTGGGAAA	8847
Qy	522	AAITCAAACTGCACCTCCGCTGACCTGGCAGAGAAAAATAGATGAGACCCCTTCAAAAGACTCC	581
Db	8848	AAITCAAACTGCACCTCCGCTGACCTGGCAGAGAAAAATAGATGAGACCCCTTCAAAAGACTCC	8907
Qy	582	AGAACTTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCA	641
Db	8908	AGAACTTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCA	8967
Qy	642	AGGATCTCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAGATCACTCCGAGA	701
Db	8968	AGGATCTCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAGATCACTCCGAGA	9027
Qy	702	AAGTCAAGGCACTTCGAGGAGAAAATTCGCCCTCTGAAAGAGAAACGTCGAGCCACGCTCAATG	761
Db	9028	AAGTCAAGGCACTTCGAGGAGAAAATTCGCCCTCTGAAAGAGAAACGTCGAGCCACGCTCAATG	9087
Qy	762	ACCTTGCTGGCAGCTTACCACTTTGGCATTTCAGCTCTCACCGTATTAACCTCAGCACTC	821
Db	9088	ACCTTGCTGGCAGCTTACCACTTTGGCATTTCAGCTCTCACCGTATTAACCTCAGCACTC	9147
Qy	822	TGGAAGACCTCAACACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAGGC	881
Db	9148	TGGAAGACCTCAACACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAGGC	9207
Qy	882	AGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGCATTTCTTTTCCACGCTCG	941
Db	9208	AGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGCATTTCTTTTCCACGCTCG	9267
Qy	942	TCCAGGCTCCCTGGGAGAGAGCCCATCTGCCCAACAAAGTGCCCTACTATATCAACACAGC	1001
Db	9268	TCCAGGCTCCCTGGGAGAGAGCCCATCTGCCCAACAAAGTGCCCTACTATATCAACACAGC	9327
Qy	1002	AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTTACAGCTTTTAGCTG	1061
Db	9328	AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTTACAGCTTTTAGCTG	9387
Qy	1062	ACCTGTAATAAGTCAGATTCTCAGCTTTAGGACTGCATGAAGTCCGAGACTCGAGA	1121
Db	9388	ACCTGTAATAAGTCAGATTCTCAGCTTTAGGACTGCATGAAGTCCGAGACTCGAGA	9447
Qy	1122	AGGCCCTTTTGTCTGGATCTCTTGGACCTGTGAGCTGTGATGCTTGGACCCAGCACA	1181
Db	9448	AGGCCCTTTTGTCTGGATCTCTTGGACCTGTGAGCTGTGATGCTTGGACCCAGCACA	9507
Qy	1182	ACCTCAAGCAAAATGACAGCCCATGATATCTCGAGATTAATTTGTTGACCACTA	1241
Db	9508	ACCTCAAGCAAAATGACAGCCCATGATATCTCGAGATTAATTTGTTGACCACTA	9567
Qy	1242	TTTATGACCGCTCGAGCAAGAGCACAACAATTTGGTCAACGCTCCCTCTCTCGCTGGATA	1301
Db	9568	TTTATGACCGCTCGAGCAAGAGCACAACAATTTGGTCAACGCTCCCTCTCTCGCTGGATA	9627
Qy	1302	TGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGTCC	1361
Db	9628	TGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGTCC	9687

QY	1362	TGTCCTTTTAAAC	CHGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGAT	1421
Db	9688	TGTCCTTTTAAAC	TGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGAT	9747
QY	1422	ACCTTTTCAAGCAAGTGGCAAGTTC	CAACAGGATTTTGTGACCAAGCGCAGGCTGGGCTCC	1481
Db	9748	ACCTTTTCAAGCAAGTGGCAAGTTC	CAACAGGATTTTGTGACCAAGCGCAGGCTGGGCTCC	9807
QY	1482	TTCTGCATGATTCATCCAA	1501	
Db	9808	TTCTGCATGATTCATCCAA	9827	
RESULT 9				
AX409637				
LOCUS	AX409637	13957 bp	DNA	linear
DEFINITION	Sequence 2284 from Patent WO229103.			
ACCESSION	AX409637			
VERSION	AX409637.1			
KEYWORDS	GI:21442342			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1			
TITLE	Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.			
JOURNAL	Gene expression profiles in liver cancer			
FEATURES	Patent: WO 0229103-A 2284 11-APR-2002;			
source	GENE LOGIC INC (US)			
	Location/Qualifiers			
	1..13957			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
	/note="EMBL/GenBank Accession No. M18533"			
ORIGIN				
Query Match	88.5%; Score 1328.8; DB 6; Length 13957;			
Best Local Similarity	99.5%; Pred.No. 0;			
Matches 133;	Conservative 0; Mismatches 7; Indels 0; Gaps 0;			
QY	162	GGGTCTTTTACAAGACAGTCTCTGACCAGTGGAGCGTCTGCACCTTCTCTGCAGGAAC	221	
Db	8598	GGTCCCATTTTGAAGCCAGTCTTGACAGTGGAGCGTCTGCACCTTCTCTGCAGGAAC	8557	
QY	222	TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAAGCCGACGACCTATTGGAGCG	281	
Db	8658	TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAAGCCGACGACCTATTGGAGCG	8717	
QY	282	ACTTTCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAAA	341	
Db	8718	ACTTTCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAAA	8777	
QY	342	CTAAGAACCTGTAACTCATGAGTACTCTTGAGACTGTACGAATATTCTTCACAGAGCAGC	401	
Db	8778	CTAAGAACCTGTAACTCATGAGTACTCTTGAGACTGTACGAATATTCTTCACAGAGCAGC	8837	
QY	402	CTTTTGAAGGACTAGAGAAACTCTTACAGAGAGCCACAGAGAGCTGCCTCTGAGAGAGAG	461	
Db	8838	CTTTTGAAGGACTAGAGAAACTCTTACAGAGAGCCACAGAGAGCTGCCTCTGAGAGAGAG	8897	
QY	462	CCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTTGAGTGGGAAA	521	
Db	8898	CCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTTGAGTGGGAAA	8957	
QY	522	AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCC	581	
Db	8958	AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCC	9017	
QY	582	AGGAACCTCAAGAGGCCACGGATGAGCTGGACTCAAGCTCGGCCAAGCTGAGGTGATCA	641	
Db	9018	AGGAACCTCAAGAGGCCACGGATGAGCTGGACTCAAGCTCGGCCAAGCTGAGGTGATCA	9077	

AUTHORS	Chamberlain, J.S. and Harper, S.Q.									
TITLE	Mini-dyotrophin nucleic acid and peptide sequences									
JOURNAL	Patent: WO 0229056-A 1 11-APR-2002;									
	THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)									
FEATURES	Location/Qualifiers									
source	1..13957									
	/organism="Homo sapiens"									
	/mol_type="unassigned DNA"									
	/db_xref="taxon:9606"									
ORIGIN										
Query Match	88.5%	Score	1328.8;	DB	6;	Length	13957;			
Best Local Similarity	99.5%	Pred. No.	0;							
Matches 133;	Conservative	0;	Mismatches	7;	Indels	0;	Gaps	0;		
QY	162	GGGTCTCTTTTCAAGACAGAGTCTTGACACAGTGGAGAGCGTCTGCACCTTTCTCTCGAGAAC	221							
DB	8598	GGTCCCATTTGGAAGCCAGTTCTGACCACTGGAAGCGCTCTGCACCTTCTCTCGAGAAC	8657							
QY	222	TTCTGGTGTGGCTACAGCTCAAGAGATGATGAATTTAGCCGCGAGGACACCTATTGGAGCG	281							
DB	8658	TTCTGGTGTGGCTACAGCTCAAGAGATGATGAATTTAGCCGCGAGGACACCTATTGGAGCG	8717							
QY	282	ACTTTCCAGCAGTTTCAGAACAGAAACGATGTACATAGGGCTTCACAGAGGGAATTCAGAAA	341							
DB	8718	ACTTTCCAGCAGTTTCAGAACAGAAACGATGTACATAGGGCTTCACAGAGGGAATTCAGAAA	8777							
QY	342	CTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGC	401							
DB	8778	CTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGC	8837							
QY	402	CTTTTGAAGACACTAGAGAAACTCTTACCAGAGAGCCACAGAGAGCTGCCTCTGAGGAGAG	461							
DB	8838	CTTTTGAAGACACTAGAGAAACTCTTACCAGAGAGCCACAGAGAGCTGCCTCTGAGGAGAG	8897							
QY	462	CCCAGAAATGTCACTCGGCTTTCTACGAAAGCAGGCTGAGGAGTCAATATCTGAGTGGGAAA	521							
DB	8898	CCCAGAAATGTCACTCGGCTTTCTACGAAAGCAGGCTGAGGAGTCAATATCTGAGTGGGAAA	8957							
QY	522	AATTGAACCTGCATCCGCTGACTGGCAGAGAGAAAATAGATGAGACCCCTTGAAGACTCC	581							
DB	8958	AATTGAACCTGCATCCGCTGACTGGCAGAGAGAAAATAGATGAGACCCCTTGAAGACTCC	9017							
QY	582	AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCAAAGCTGAGTGATCA	641							
DB	9018	AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCAAAGCTGAGTGATCA	9077							
QY	642	AGGATCTCGCAGCCGCTGGCGATCTCTCTCATTGACTCTCTCCAAAGATCACTCGAGA	701							
DB	9078	AGGATCTCGCAGCCGCTGGCGATCTCTCTCATTGACTCTCTCTCCAAAGATCACTCGAGA	9137							
QY	702	AAGTCAAGGCACATTCGAGGAGAAATTCGGCTCTGAAAGAGAACGTCGAGCCACGCTCAATG	761							
DB	9138	AAGTCAAGGCACATTCGAGGAGAAATTCGGCTCTGAAAGAGAACGTCGAGCCACGCTCAATG	9197							
QY	762	ACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAAAGCTCAGCACTC	821							
DB	9198	ACCTTGCTCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAAAGCTCAGCACTC	9257							
QY	822	TGGAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAAGC	881							
DB	9258	TGGAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAAGC	9317							
QY	882	AGCTGTCATGAAGCCACACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGCTG	941							
DB	9318	AGCTGTCATGAAGCCACACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGCTG	9377							
QY	942	TCCAGGCTCCCTGGGAGAGACCATCTCGCCAAACAAAGTGCCTACTATATCAACCAAG	1001							
DB	9378	TCCAGGCTCCCTGGGAGAGACCATCTCGCCAAACAAAGTGCCTACTATATCAACCAAG	9437							
QY	1002	AGACTCAAAACAACTTGCTGGGACCATCCCAAAATTCAGAGAGCTCTACAGTCTTTAGCTG	1061							

in different patients were from positions 302-2200, 473-1168, 1691-1810, and 1169-3011.

FEATURES	Source
Location/Qualifiers	
1..13957	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/chromosome="X"
	/map="Xp21.3-p21.1"
	/tissue_type="muscle"
	/dev_stage="fetus"
	/note="G00-119-850"
1..13957	
/gene="DMD"	
209..11266	
/gene="DMD"	
/codon_start=1	
/product="dystrophin"	
/protein_id="AA053189.1"	
/db_xref="GI:181857"	
ODGRLDLDLLEGLTGOKLPKEKSTRVLAHNVNKAIRLVQNNVDLVNIGSDIDVDG	
NKLTLGLIWNIIILHWQVKNVMKNI MAGLOQTNSEKILLSWVRQSTPNVQVNIKFT	
TSWSDGLALNALIHSHRDPDLDNWSVVCQASATQRLPHEANILARYQLGIGKLDPDV	
DITTPDKSILMYITSLFQVLPOQVSTEAICQVEMLRPPKVKKEBHFQHHQWHTSQ	
QITVSLAQGERTSSPRFKASYIAQAAVITSDPTRSPFQSHLEAPDKFGFSSLL	
MESVNIIDRYQTALIEEVLWSLLSAEDTLQAGEISNDVSVKQDFHTHEGYMMDLTAH	
QGRVNTLOIGSKLIGTGLKSEDETEVQGMNLLNSRWECLRVASKEGNSMLRLVIM	
DLQOKLKEINLMDLTBEETRKMEEPPLGDLDLKROVQCHVKVLEDELEOBRVIM	
SLTHVWNPDESSDHTAALEOLKVLGDRWANI CRWTEDEWLVLLQDILLKWRLTE	
EQCLFSAWLSEKEDAVNKIHTGFKQDNELSSIQLKVLKADLEKKQSMGKLSLK	
QDLISTLKNKSVTKTEAMLDNFARCMNDLVQKLEKSTAQISQAVITITPSLTQTIVM	
ETVITTVITFOILVHAQEEPLPPPPQKKRQITVDSEIRKRLDVIDTELHSWITRSEA	
VLOQPEFAIRKESGNDLKEKVNAIERKAEFRKLQDASRAQALVEQVWNGVNA	
DSIIKQAGEQNSRWIEFQCLSERLWLEYQNNIIIPYNQLOQLEQMTTAAENWLKIQ	
PTTSEPTAQAQSEKIKI CKDEVNLSLGHQPOIERLKIOSIALKEKGQPMFLDADVFV	
TNHPFQVSDVQAREKLOTTIDTLPMRYQETMMSAIRTWVQOSETKLSTPOLSVTDY	
ELMEORLEQALQSSLOQSGLYLSTTVKEMSKAPSEISRKQSEFEIEGRWK	
KLSSQLVEHQQLQEEQNSKURKIQNHILTKKMAEVDVFLKESEFALGDSLELKQQL	
KQSLVDSDITQIPSNVNEGGQIKNEAPEFASRLTELKELNLTQWDMHCQVY	
ARKGALGKGLVSLQDLSSEHWMVTOAEYELDERDFEYKTPDELOKAVEEMKRAK	
EAAQKEAKVLTESVNSVIAQAPVPAQALKLELTITNTQMLCTRNLNGCKTLE	
EVAACHELLISLEFANKMLNEVEFKLTNTENPGGAEEISEVLDLSLENMRSEBNP	
NOIRLAQITDGGWMDLNEBELTENSREBELHEAVROKLILOSIOQAQETSKS	
LHLQESLITFDIKQALAYADKVDAAQMPQEAQKIQSDLSISHEISEMKKNQKGBA	
AQRVSLQDVAKKQLQVSMKRFUQKPFANFELQKSKMILDEVKNHJLPALSTKSV	
QEVVQSNHCNVNLYKSLSEVKEVEMVINTGRQIVQKQTNENKELDERVTALKLHY	
NELGAKVTERKQLEKCLKSRMRKEMVLTWLAATDMELTKRSAGEVMPNLDSE	
VAMGKATQKIEKQVHLKSTIEVGEALKTVLGKKTLLVEDKLSLNSNWI AVTSRAE	
EWMLLLEYOKHMETPONVDHITKIIQADTLIDSEKKKPOOKEDVLRKLABEIND	
IRPKVDSTRQAANLMNRGDHCKLVEPOISLNRHPRALSHRIKTKGASIPLEKELE	
QPNSDIKLLEPLEAEITQQGVNLKEEDFNKMDNEGTVKELLQGDNLQQRITDER	
KREBIKIQQLQTKGNALDKLRSORRKALEISHQWYQYKQADLLKCLDDIEKKL	
IASPEPRDERKIEIDELQKKEELNVRRAQEGSEDAVAVPTQTLQSKWRRE	
TESFAQFRRLNFAQIHTVRETMVMVMDPLEISVYPSTYLTETHTVSQALLEYVQ	
LLNAPDIKADFDLFOEESLAKIKDSLOQSSGRIDIHSKKTALQATPVREYKL	
QEALSOLDFOEWKVMYKDRQGFDRSVKRRFRHYDIKIPNQMTEABQFRTKQI	
PENWEHAKYKWLKELODGIQQRQTVRTLNAGEBIIQOSSKTASIIQEKLSLNL	
RWQVCHQSDRKRRKEQNNISEFQDNEFVLMDEADNATASIPPEPKEQQLKE	
KLEQVLLSELPLEPRLQIKLQNETGGPVLVSAPISEEQDKLENKLKQTNQWIKVS	
RALPEKGEIEAQIKDLQLEKLEDELEQHLHLLWSPIRNQLKIIYNQNOEGFVD	
VQETEIAVQAKQDPVEBISKGQHLVYKEKPAQPVKRLKEDLSSEWKNVRLIQELRA	
KQPLAPRLNTF IGASPTQTVTLVTPVPTVETAKISLEMPSSILMVPALADFNRAWT	
KIDTLWLSLQVIVKQSRVMDLEDINEMLIKOKATWQDLEQRRPOLBELITAAONLK	
NKTNSQPARTIIDRIBRIQNWDEVEHJONRRQQLNEMLKDSTOWLEKKEAEQVL	
GQAKLESKEGPIYVDIAIQKKITETQJAKDLRQWQTNVDVANDLAKLRDYSAD	
DTRVKNMITENINASWRSIHKRSVEREAALETHRLQFPDLDFLAWLFEATTA	
NVLDQATRKERLLEDXGKVELAKMQODLOQIEEAHTDVYVNLNDENSQKILRLSEGD	
DAVLLQRLDNMNFKWSLRKKSINIRSHLEASSDQWKRHLHLSIQLLWVWLQKLDQ	
SROAPIGDPTPAVQKNDVHRAFKRLKTKPEVIMSTMLETVRIFLTEOPLGLEKLYQ	
EPRELPEERAVQNTVRLLRQAQSEVNTENEKLNHSDWQKIDETLERLOELQEAR	
ELDKLQRAEVPNLSTLEDINLRWKLQVAVEDVRVQLHEAHRDFGASQPLSTSVQGP	

ORIGIN	Query Match	88.5%	Score 1328.8	DB 9	Length 13957
	Best Local Similarity	99.5%	Pred. No. 0		
	Matches 1333	Conservative	0	Mismatches	7
				Indels	0
				Gaps	0
QY	162	GGGTTCTTTTACAAGACAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTTGCGAGGAAC	221		
Db	8598	GGTCCCAATTGGAAGCCAGTCTGACCAGTGAAGCGTCTGCACCTTTCTTGCGAGGAAC	8657		
QY	222	TTCTGCTGTGGCTACAGCTGAAGATGATGAATTAAGCGCGCAGCAGCCTATTGGAGGGG	281		
Db	8658	TTCTGCTGTGGCTACAGCTGAAGATGATGAATTAAGCGCGCAGCAGCCTATTGGAGGGG	8717		
QY	282	ACTTTCCAGCAGTTCAGAACGAGACGATGTATCATAGGCGCTTCAAGAGGGGAATTAATA	341		
Db	8718	ACTTTCCAGCAGTTCAGAACGAGACGATGTATCATAGGCGCTTCAAGAGGGGAATTAATA	8777		
QY	342	CTAAGAAGCCTTAATCATGACTTCTTGAGACTGTACGATATTTCTGACAGAGCAGC	401		
Db	8778	CTAAGAAGCCTTAATCATGACTTCTTGAGACTGTACGATATTTCTGACAGAGCAGC	8837		
QY	402	CTTTGGAAGGACTAGAGAACTCTTACCAGAGCCAGAGAGCTGCTCTCTGAGGAGAGAG	461		
Db	8838	CTTTGGAAGGACTAGAGAACTCTTACCAGAGCCAGAGAGCTGCTCTCTGAGGAGAGAG	8897		
QY	462	CCCAGATGCTACTGGCTTCTACGAAAGCAGCGCTGAGAGGTCAATCTAGTGGGAAA	521		
Db	8898	CCCAGATGCTACTGGCTTCTACGAAAGCAGCGCTGAGAGGTCAATCTAGTGGGAAA	8957		
QY	522	AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC	581		
Db	8958	AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC	9017		
QY	582	AGGAATCTCAAGAGGCCACGATGAGCTGAGCCTCAAGCTGCGCAAGCTGAGGTGATCA	641		
Db	9018	AGGAATCTCAAGAGGCCACGATGAGCTGAGCCTCAAGCTGCGCAAGCTGAGGTGATCA	9077		
QY	642	AGGGATCTCGCAGCGCGTGGCGGATCTCTCATTTGACTCTCTCCAAAGATCACCTCGAGA	701		
Db	9078	AGGGATCTCGCAGCGCGTGGCGGATCTCTCATTTGACTCTCTCCAAAGATCACCTCGAGA	9137		
QY	702	AAAGTCAAGGCACTTCGAGAGAGAAAATTGGCGCTCTGAAAGAGAACGTCAGCCACGTCATG	761		
Db	9138	AAAGTCAAGGCACTTCGAGAGAGAAAATTGGCGCTCTGAAAGAGAACGTCAGCCACGTCATG	9197		
QY	762	ACCTTGCTCGCAGCTTACCACITTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGACTC	821		
Db	9198	ACCTTGCTCGCAGCTTACCACITTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGACTC	9257		
QY	822	TGGAAGACCTGAAACCAAGATGGAAGCTTCTGACGCTGGCGCTCGAGGACCGAGTCAGGC	881		
Db	9258	TGGAAGACCTGAAACCAAGATGGAAGCTTCTGACGCTGGCGCTCGAGGACCGAGTCAGGC	9317		
QY	882	AGCTGATGAAGCCCAAGGAGACTTTGGTCAGCATCTCAGACATTTCTTCCACGCTG	941		
Db	9318	AGCTGATGAAGCCCAAGGAGACTTTGGTCAGCATCTCAGACATTTCTTCCACGCTG	9377		
QY	942	TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAAGTGCCTACTATATCAACACAG	1001		
Db	9378	TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAAGTGCCTACTATATCAACACAG	9437		

```
QY 1002 AGACTCAAACTTGTGGGACCATCCCAATATGACAGAGCTCTACAGTCTTTAGCTG 1061
Db 9438 AGACTCAAACTTGTGGGACCATCCCAATATGACAGAGCTCTACAGTCTTTAGCTG 9497
QY 1062 ACCTGAATAATGTGACAGATTTCTCAGCTTATAGGAGTGCATGAACTCCGGAAGACTGCAGA 1121
Db 9498 ACCTGAATAATGTGACAGATTTCTCAGCTTATAGGAGTGCATGAACTCCGGAAGACTGCAGA 9557
QY 1122 AGGCCCTTTGTTGGATCTCTTGAGCTGTGACGTGATGATGCTTGGACAGACACA 1181
Db 9558 AGGCCCTTTGTTGGATCTCTTGAGCTGTGACGTGATGATGCTTGGACAGACACA 9617
QY 1182 ACCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATTAATTTGTTGACCACTA 1241
Db 9618 ACCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATTAATTTGTTGACCACTA 9677
QY 1242 TTTATGACCGCTGGAGCAAGAGACACAACAATTTGGTCAACGCTCTCTGCGTGATA 1301
Db 9678 TTTATGACCGCTGGAGCAAGAGACACAACAATTTGGTCAACGCTCTCTGCGTGATA 9737
QY 1302 TGTGCTCAACTGCTGCTGATGATTTTATGATGCGGACGACAGAGGAGGATCCGTGTC 1361
Db 9738 TGTGCTCAACTGCTGCTGATGATTTTATGATGCGGACGACAGAGGAGGATCCGTGTC 9797
QY 1362 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGACAAGTACAGAT 1421
Db 9798 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGACAAGTACAGAT 9857
QY 1422 ACCTTTTCAACAGTGGCAAGTTTCAACAGATTTTGTGACCGGACGAGCTGGGCTCC 1481
Db 9858 ACCTTTTCAACAGTGGCAAGTTTCAACAGATTTTGTGACCGGACGAGCTGGGCTCC 9917
QY 1482 TTCTGCAATGTTCTATCCAA 1501
Db 9918 TTCTGCAATGTTCTATCCAA 9937

RESULT 12
LOCUS AR220819 13977 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 60 from patent US 6426186.
ACCESSION AR220819
VERSION AR220819.1 GI:2327696
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 13977)
AUTHORS Jones,K.A., Volkunth,W. and Walker,M.G.
TITLE Bone remodeling genes
JOURNAL Patent: US 6426186-A 60 30-JUL-2002;
FEATURES
Location/Qualifiers
1..13977
/organism="unknown"
/mo_type="genomic DNA"

ORIGIN
Query Match 87.8%; Score 1317.8; DB 6; Length 13977;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 162 GGGTCTTTTACAGACAGTCTGACCAAGTGGAGCGTCTGACCTTTCTCTGAGGAAAC 221
Db 8598 GGTCCATTTTGAAGCCAGTTCTGACCAAGTGGAGCGTCTGACCTTTCTCTGAGGAAAC 8657
QY 222 TTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGACGACCTATTGGAGCG 281
Db 8658 TTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGACGACCTATTGGAGCG 8717
QY 282 ACTTTCCAGCAGTTTCAAGCAGACGATGTACATAGGCTTCAAGAGGGAATTTGAAA 341
Db 8718 ACTTTCCAGCAGTTTCAAGCAGACGATGTACATAGGCTTCAAGAGGGAATTTGAAA 8777
```

```
QY 342 CTAAAGAACTGTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 401
Db 8778 CTAAAGAACTGTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 8837
QY 402 CTTTGGAAAGACTAGAGAACTCTACAGAGAGCCAGAGAGCTCCCTCTTGAGGAGAGAG 461
Db 8838 CTTTGGAAAGACTAGAGAACTCTACAGAGAGCCAGAGAGCTCCCTCTTGAGGAGAGAG 8897
QY 462 CCCAGAACTGTCACTCGGCTTCTAGAAAGCAGAGCTGAGGAGTCAATACTGAGTGGGAAA 521
Db 8898 CCCAGAACTGTCACTCGGCTTCTAGAAAGCAGAGCTGAGGAGTCAATACTGAGTGGGAAA 8957
QY 522 AATTGAACTCCTGACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGATCC 581
Db 8958 AATTGAACTCCTGACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGATCC 9017
QY 582 AGGAACTTCAAGAGAGCCAGGATGAGCTGAGACCTCAAGCTGCGCAAGCTGAGTGATCA 641
Db 9018 AGGAACTTCAAGAGAGCCAGGATGAGCTGAGACCTCAAGCTGCGCAAGCTGAGTGATCA 9077
QY 642 AGGGATCTCTGCGAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAAGATCACCTCGAGA 701
Db 9078 AGGGATCTCTGCGAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAAGATCACCTCGAGA 9137
QY 702 AAGTCAAGGCACTTGGAGGAGAAATTTGGCGCTCTGAAAGAGAAAGTGGAGCCACGCTCAATG 761
Db 9138 AAGTCAAGGCACTTGGAGGAGAAATTTGGCGCTCTGAAAGAGAAAGTGGAGCCACGCTCAATG 9197
QY 762 ACCTTGCTGCGCAGCTTACCACTTTGGGCACTTCAAGCTCTCAAGCTTAAACCTCAGACACTC 821
Db 9198 ACCTTGCTGCGCAGCTTACCACTTTGGGCACTTCAAGCTCTCAAGCTTAAACCTCAGACACTC 9257
QY 822 TGGAAAGACTGAACAACAGATGGAGGTTCTGAGGTGGCGCTCGAGAGACGAGCTCAGGC 881
Db 9258 TGGAAAGACTGAACAACAGATGGAGGTTCTGAGGTGGCGCTCGAGAGACGAGCTCAGGC 9317
QY 882 AGCTGCATGAAGCCACAGGAGCTTTGGTCCAGATCTCAGACCTTCTTTTCCAGCTCTG 941
Db 9318 AGCTGCATGAAGCCACAGGAGCTTTGGTCCAGATCTCAGACCTTCTTTTCCAGCTCTG 9377
QY 942 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTCCCTACTATATCAACCAAC 1001
Db 9378 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTCCCTACTATATCAACCAAC 9437
QY 1002 AGACTCAAACTTGTGAGACCATCCCAATATGACAGAGCTCTACAGTCTTTAGCTG 1061
Db 9438 AGACTCAAACTTGTGAGACCATCCCAATATGACAGAGCTCTACAGTCTTTAGCTG 9497
QY 1062 ACCTGAATAATGTGACAGATTTCTCAGCTTATAGGAGTGCATGAACTCCGGAAGACTGCAGA 1121
Db 9498 ACCTGAATAATGTGACAGATTTCTCAGCTTATAGGAGTGCATGAACTCCGGAAGACTGCAGA 9557
QY 1122 AGGCCCTTTGTTGGATCTCTTGAGCTGTGACGTGATGATGCTTGGACAGACACA 1181
Db 9558 AGGCCCTTTGTTGGATCTCTTGAGCTGTGACGTGATGATGCTTGGACAGACACA 9617
QY 1182 ACCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATTAATTTGTTGACCACTA 1241
Db 9618 ACCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATTAATTTGTTGACCACTA 9677
QY 1242 TTTATGACCGCTGGAGCAAGAGACACAACAATTTGGTCAACGCTCTCTGCGTGATA 1301
Db 9678 TTTATGACCGCTGGAGCAAGAGACACAACAATTTGGTCAACGCTCTCTGCGTGATA 9737
QY 1302 TGTGCTCAACTGCTGCTGATGATTTTATGATGCGGACGACAGAGGAGGATCCGTGTC 1361
Db 9738 TGTGCTCAACTGCTGCTGATGATTTTATGATGCGGACGACAGAGGAGGATCCGTGTC 9797
QY 1362 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGACAAGTACAGAT 1421
Db 9798 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGACAAGTACAGAT 9857
QY 1422 ACCTTTTCAACAGTGGCAAGTTTCAACAGATTTTGTGACCGGACGAGCTGGGCTCC 1480
```

Db 9858 ACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCAAGCGCAGCTGGGCGCTC 9917
QY 1481 CTCTCTCATGATCTATCAAA 1501
Db 9918 CTCTCTCATGATCTATCAAA 9938
RESULT 13
AF070485
LOCUS 13887 bp mRNA linear MAM 09-DEC-1998
DEFINITION Canis familiaris dystrophin mRNA, complete cds.
ACCESSION AF070485
VERSION AF070485.1 GI:3982750
KEYWORDS
ORGANISM Canis familiaris (dog)
Canis familiaris
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 13887)
Carville, K.S., Mann, C.J., Scharzberg, S.J. and Wilton, S.D.
Direct Submission
Submitted (04-JUN-1998) ANRI, Pathology, University of Western
Australia, Verdun Street, Nedlands, WA 6018, Australia
Location/Qualifiers
1..13887
/organism="Canis familiaris"
/mol_type="mRNA"
/strain="breed Golden retriever"
/db_xref="taxon:9615"
<1..289
290..11332
/codon_start=1
/product="dystrophin"
/protein_id="AAC83646.1"
/db_xref="GI:3982751"
/translation="MLWEEVEVCYREDVQKKTFTKWNNAOFKFKOHENLPSDL
QDGRLLDLEGITGQGLPREKSTVRVHALNNVKNALRYLQKNVLDNIGSTDIVD
NHLTLGLNIILHVNQVKNMKNIMAGLQNTSEKILLSWRQSTNRYPNQVNI
TSWDLGLNALNLSHRPDLFVNSVVCOSQATRLEHAFNIKAYQTEKLLDPEDV
ATYPDKSLMIVTSLFQVLPQVSEIAIOEVLPRPQVTRREHFOIHHMVYSQ
QITVSLAQGVTERAPKPKFATYATQAAVYVITFTSPLPSQHLETPEKSFORS
LITEANLDSYQALFEVLSWLLSADALQAQGEISNDVEEVEKQETHHEGNMUTS
HQRQVNLGLNMLTKTEREKEKEPLGPDIEDLKQVQKHVLEDEQEVRY
MLQNLQGLNMLTKTEREKEKEPLGPDIEDLKQVQKHVLEDEQEVRY
NSLTHMVVYDESSGHATAALEQKLVLDWANTCRWEDRWLLQILKWKQRT
BECCLFSAKLSXEDAVNKHHTTGKQDSEVLNLOKLVLTDLKKQTDMDKCLSL
NQDLISALNTVVAHMEAWLNDFAQRMDNLVQKLSKSNQISOAVTTTQPSLTQTV
METVTMTTTHLILVKAQELPPPPKQKQLIVSEIRKLDVDTIELHSHWIRSE
AVLOSPEFATYRKEGNSDLKEKNAIEREKAERKQDASRAQALVEQWVNEGNT
ADSIKQSEQLNRSWIEFQGLSERLNWLEQNNIITFYNLQQLQEQMTTAEENLKT
QPTTTSPTAIKSKICKEINRLSALQPIERLKIQTIALKEKQGFMDADFA
FTNHNQVADVAAREKELOTIFDLSPPMRYQETMTILTIIQOSETKLSIPQVITE
YDMEQRLGELQALSSLOQNGNLVLTIVKEMSKAPLSDISRKYOSEPEIEGR
WKULSOLVHECQKLEQMAKLRKIQHNIKTLKWITEVDVFLKEWPALEILKR
QLKQRLVINDIQTIPSLNVEGAQKNNEAFEPFAGRELTELRELTNTQWDMYCRQ
VIARKEALKGSLDLSQKLDSEHMTQAEVYLERDLEFYTDELOTAVEENKR
AKEBAQKQKELVLSVNSVIAQAPPAQEAALKEKELDTLTNYQLWCLTRLNGCKT
LEEWNAQWHELLSVEKANKWSEVVKLKTENISGAEIEAELVLSLENLMQSHED
NPQNLRIALQITDGGVDELINDELFTNSRWRELHEAVRQKLEOSIQOAGIE
KSLHLIQESLSSIDQLAAVITADKVDAAQMPQEAQKIQSDLTSHESLEEMKHNOG
ETARQVLSIDVQKLDQVSMKFLQKPAQNTPEQRLQESKMLDVEKHLPALETKS
VEQVVSQNLHNLKYSLEVKSEVEMVKTGRQIVQKKQNTENPELDERVATLKL
HYNELGAKVTERKQLEKCLKSRMRKEMNALTWEAATDMELTKRSVEGMPSLD
SEVANGATQKEIEKQVHLKSVTEGALKTLVLGKEMELVKLSDNLSNIAVTSR
AEBWNLLEYQKHMTFDQNVDTIWNIIQADALLDESEKKKPOQKEDLKLKXEM
NDIRPKVDSFRDQANIMANRCHRCVVRPKISELNRFAAISHRIKGTGASIPLKE
LEQNSIQKLEPLEAIIQGVNLKEEDNECTVKELQVGNLQIRGDNLOQITD
ERKEEIKIQKLLQTHNALKOLRSRRKALIEISHQVYQKQADDLKLCDDIEK
KLASLPFRERKIEIDRELQKXEBELNAVROAEGLSDEGAMAVEPTQIQLSRW
REIESKAFQRLNFAQHTVHEESVAMTDEMPLEISYVPTYLEITHTVSQALSE
BELLNAPDLCAQDEDFKQESLKNIKDLSQISGRIDIHNKKTAAALSPAPRA
KLQALSRLLDQWERNVNNYKDRQGRFDRSVEKWRFRHYDMKILNQLWLEAEQFLKKT

QIPNWEHAKYKWLKELODGIGROSIVRVLNATGEEIIQOSKFLTSDASILQKLGSL
NLWQEVCKQLAERKKELEKQNLSEFORDVNEFVLMLEADNVANIPIEPGNEQOL
KELEQKLLAEELPLRQGLKQLNETGTVLSAPLSPEDQKLENKLTQTNLOLK
VVRNLPKQBEIEBAHVKDLQGLBEQNLHLLWLSPIRNLQENINQPNQTFDKEIE
VAVQAKQPDVEGILSKGHLYKPKPATQPAKRLLEDLSWKVVTQLIQERAKQGP
APGLTTRVAPPSQTVLTQPAVTKETAIKLEMPSSLLLEPALDFNFAWTELTDW
LSLLDRVIKSORVMVGLDENIMI IKQATLQDLQORRQLEELITAAQNLKNTSN
QEARTIITRIEIOSOWDEVOHQLNRQLTQMLKQSTQWLEAKEAEQVGLQARA
KLESKEAPYTVDAIOKKTETQKLDLQKQINVDVANDLAKILRDYSADTRKV
HMITENINASWASHIKLSERBAALETHRLLOQFLDLEKELATEETANVLOD
ATHERLLEDKGVRELMKQWDLQGEIEARTDIYHNLDSNGKVLASLSDDAALL
ORRLDNVNFWSERLKKSLNIRSHLEASSQWRLHLSLOELVWLQKODELSQAP
IQGFPAVQKQNDVHRAFEKRLTKFVIMSTLETETFLTEFQLEGLKLYQBPRL
PPBERAQNVTRLRQAEVNTQWELKNSHSDWQRKIDLEALRQLEQATDELDK
LROAEVIGKSPQVGLDILSDLOADEKVKALGEITPLKENVSYNDLARQTLTGLI
LQSPYNLTLEDINTNPKLLQVADIEDRIOLHEARDFGPASOHLFVLSVQGPWERAI
SPNKPVIINHTQTTCTDHEPKTELYQSLADLNNVPSAYRTAKLRRLOKLCILD
LSLSAADCALDQHNKQNDQMDLQVINCLETIYDRLEHNNLNVPLCYDMCLNWK
LLNYDQRTGRIRVLSFKTGISLKAHLEDKIRYLFQKVASSTQCDORRGLLH
DSIQIPRLQGEVAFSGNSIETPSVSCFPANNKPEIEAALFLDMKLEFQSMWMLPV
LHRVAAETAKHQAKNICKEPTIGFRYSLKHFNVDIQCSCFFSRVAKGHVYP
MVEYCTPTSGEDVRDPKVLKNKFTKRYFAKPRMGYLQVOTVLEGNMETPVTLI
NFWFVDSAPASSPOLSHDDTHSRLEHYASRLKNVNSGSVYNDISPNISIDDEHLL
IQHWRSNQESPLSPQSPRAQIILISSEERLELILADLEGRNLSLOAEVDRKQ
QHEHGLSPSPSPPMPTSPQSPRAELIAAKLRQKGRLEARNQILEDHKNQLE
SQLHLRQLLEQPOAEAKVNGTTVSPSTSLQSDSSQPMLLRVVGSQTSSEMEEDL
LSPPDFTSTGLEEVMEQLNHSFSSRGRTPEKPMREDTM"
11333..>13887
3'UTR
repeat_region
/rpt_type=tandem
/rpt_unit="ca"
ORIGIN
Query Match 79.0%; Score 1185.4; DB 4; Length 13887;
Best Local Similarity 92.8%; Pred. No. 0;
Matches 1243; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
QY 162 GGGTCTTTTACAAGACAGTCTGACAGTGGAGCGTCTGCACCTTCTCTGAGGAGAC 221
Db 8664 GGTCTCACTTGGAAACCCAGTCTGACAGTGGAGCGTCTGCACCTTCTCTGAGGAGAC 8723
QY 222 TTCTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGCGCGGAGCAGCCTATTGGAGCG 281
Db 8724 TTCTGCTGTGGCTTCCAGCTGAAAGATGATGAATTAAGCGCGGAGCAGCCTATTGGAGCG 8783
QY 282 ACTTTCAGCAGCTTCCAGAGCAGAACGATGATACATAGGCGCTTCAAGAGGGAATTGAAA 341
Db 8784 ACTTTCAGCAGCTGCGAGAGCAGAACGATGATACATAGGCGCTTCAAGAGGGAATTGAAA 8843
QY 342 CTAAGAAGACCTTAATCATGATCTCTGAGACTGTACGAATATTTCTGACAGAGCAGC 401
Db 8844 CGAAAGAACCTTAATCATGATCTCTGAGACTGTACGAATATTTCTGACAGAGCAGC 8903
QY 402 CTTTGGAAAGACTAGAGAAACTCTTACCAGGAGCCAGAGAGCTGCTCTCTGAGGAGAG 461
Db 8904 CTTTGAAGGACTAGAGAACTCTTACCAGGAGCCAGAGAGCTGCTCTCTGAGGAGAG 8963
QY 462 CCCAGAAATGCTACTCGGCTTCTACAAAGCAGGCTGAGAGGTCATTAATGAGTGGGAAA 521
Db 8964 CCCAGAAATGCTACACGGCTCTCTACGAAAGCAGGCTGAGAGGTCATTAATGAGTGGGAAA 9023
QY 522 AATTGAACCTGCACCTCGGCTGACTGCGCAGAGAAAAATAGATGAGACCTCTTGAAGACTCC 581
Db 9024 AACTGAACCTGCACCTGCGAGACTGCGCAGAGAAAAATAGAGAGGCTCTCGAAGACTCC 9083
QY 582 AGGAACTTCAAGAGGCCACCGATAGCTGGAGCCTCAAGCTGCGCAGCAGCTAGGTGATCA 641
Db 9084 AGGAGCTTCAAGAGCAACAGATGAGCTGGATCTCTCAAACTAGCTCAGGAGAGGTGATCA 9143
QY 642 AGGATCTTGGCAGCCGCTGGCGGATCTCTCATGACTCTCTCAAGATCACTCTCGAGA 701
Db 9144 AAGGATCTTGGCAGCCCTTGGGTGACCTCTCTCATGACTCTCTCAAGATCACTCTCGAGA 9203

QY 702 AAGTCAAGGCACTTCGAGGAGAAATTCGCTCTCTGAAAGAGAGCTGAGCCAGCTCAATG 761
Db 9204 AAGTCAAGGCGCTTCGAGGAGAAATTCACCTCTGAAAGAGAAATGTGAGCTAGCTCAATG 9263
QY 762 ACCTTGCTCGCAGCTTACCACTTTGGGCACTTACCTCTGAGCTCTACCGTATTAACCTGAGCTC 821
Db 9264 ACCTTGCTCGCAGCTTACCACTTTGGGCACTTACCTCTGAGCTCTACCGTATTAACCTGAGCTC 9323
QY 822 TGAAGAGCTGAACACCAAGATGGAAGCTTCTGCAAGTGGCGCTGAGGAGCCAGTCAAGC 881
Db 9324 TGAAGAGCTGAACACCAAGATGGAAGCTTCTGCAAGTGGCGCTGAGGAGCCAGTCAAGC 9383
QY 882 AGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTGAGCACTTTCTTTTCCAGCTCTG 941
Db 9384 AGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTGAGCACTTTCTTTTCCAGCTCTG 9443
QY 942 TCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCAAG 1001
Db 9444 TCCAGGCTCCCTGGGAGAGAGCCATCTCACCAACAAAGTGCCTACTATATCAACCAAG 9503
QY 1002 AGACTCAAACTTGTGTGGGAGCCATCCCAAAATGACAGAGCTTACCACTCTTTAGCTG 1061
Db 9504 AGACCCAAACAACTTGTGTGGGAGCCATCCCAAAATGACAGAGCTTACCACTCTTTAGCTG 9563
QY 1062 ACCTGAATATATGTCAGATTTCTGAGCTTATAGGACTGCCATGAACCTCCGAGAGCTCGA 1121
Db 9564 ACCTGAATATATGTCAGATTTCTGAGCTTATAGGACTGCCATGAACCTCCGAGAGCTCGA 9623
QY 1122 AGGCCCTTTGCTGGATCTCTTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 1181
Db 9624 AGGCCCTTTGCTGGATCTCTTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 9683
QY 1182 ACCTCAAGCAAAATGACAGCCATGATATCTGAGATTTCTGAGATTTATTTGTTGACCACTA 1241
Db 9684 ACCTCAAGCAAAATGACAGCCATGATATCTGAGATTTCTGAGATTTATTTGTTGACCACTA 9743
QY 1242 TTTATGACCCCTGGAGAGAGACCAACAAATTTGGTCAAGCTCCCTCTCTGCTGGTGGATA 1301
Db 9744 TTTATGATGCTCTAGAGAGAGAGACCAACAAATTTGGTCAAGCTCCCTCTCTGCTGGTGGATA 9803
QY 1302 TGTGCTGAAGTGGCTGCTGAATTTTATGATACGGGAGCAACAGGAGAGATCCCGTGTCC 1361
Db 9804 TGTGCTGAATGGCTGCTGAATTTTATGATACGGGAGCAACAGGAGAGATCCCGTGTCC 9863
QY 1362 TGTCTTTTAAATGCTGATCATTTCCCTGTGTAAGACACATTTGGAGAGCAAGTACAGAT 1421
Db 9864 TGTCTTTTAAATGCTGATCATTTCTCTGTGTAAAGCCATTTGGAGAGCAAGTACAGAT 9923
QY 1422 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACCAAGCCAGGCTGGGCTCC 1481
Db 9924 ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTACCAAGCCAGGCTGGGCTCC 9983
QY 1482 TCTGATGATTTCTATCCA 1500
Db 9984 TCTGATGATTTCTATCCA 10002

RESULT 14
AX306153
LOCUS AX306153 13815 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 904 from Patent WO0188188.
ACCESSION AX306153
VERSION AX306153.1 GI:17645441
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Ishikawa, K., Asai, S., Takahashi, Y., Nagata, T. and Ishii, Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 904 22-NOV-2001;
School Juridical Person Nihon University (JP)

FEATURES
source 1. 13815
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Query Match 76.6%; Score 1149.6; DB 6; Length 13815;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 1221; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
QY 162 GGGTTCTTTTCAAGACAGTTCCTGACAGTGGAAAGCGTCTGACCTTTCTTCGAGGAAC 221
Db 8589 GGTCCCATTTGGAACAAAGTTCTGACAGTGGAAAGCGTTTGCATCTTTCTTCGAGGAAC 8648
QY 222 TTCTGCTGTGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGCACCTTATTCGAGCG 281
Db 8649 TTCCTTTTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGCACCTTATTCGAGCG 8708
QY 282 ACTTTCAGCAGTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAA 341
Db 8709 ATTTCCAGCAGTTCAGAAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAA 8768
QY 342 CTAAGAACCTGTAATCATAGTACTCTTGAGACTGTACGATATTTCTGACAGAGAGC 401
Db 8769 CTAAGAACCTGTAATCATAGTACTCTTGAGACTGTGAGAAATTTCTGACAGAGAGC 8828
QY 402 CTTTGGAGAGTACTAGAAACTCTACAGAGAGCCAGAGAGTGCCTCTCTGAGGAGAG 461
Db 8829 CTTTGGAGAGTACTAGAAACTCTACAGAGAGCCAGAGAGTGCCTCTCTGAGGAGAG 8888
QY 462 CCCAGATGTCTCTGGCTTCTACGAAGCAGGCTGAGGAGTCAATPACTGAGTGGGAAA 521
Db 8889 CTCAGAAATGTCACTCGGCTCTACGAAGCAGGCTGAGGAGTCAACGCTGAATGGGACA 8948
QY 522 AATTGAACCTGACTCCGCTGACTGGCAGAGAAATAGATGAGACCTTTGAAAGACTCC 581
Db 8949 AATTGAACCTGCTGACTGGCAGAGAAATAGATGAGACCTTTGAAAGACTCC 9008
QY 582 AGAACTTCAAGAGCCACGAGTGAAGTGAAGTGCCTCAAGCTGCGCCAGTGAAGTGAATCA 641
Db 9009 AGAACTTCAAGAGCTGCGGATGAAGTGAAGTGCCTCAAGTTCGCGCCAGTGAAGTGAATCA 9068
QY 642 AGGATCTCTGCGAGCCGCTGGGCTCTCTGATGACTCTCTTCAAGATCACTCGAGA 701
Db 9069 AGGATCTCTGCGAGCCGCTGGGCTCTCTGATGACTCTCTTCAAGATCACTCGAGA 9128
QY 702 AAGTCAAGGCACTTCGAGGAGAAATTTGGCTCTCTGAAAGAGAACTGAGCCAGCTCAATG 761
Db 9129 AAGTCAAGGCACTTCGAGGAGAAATTTGCACTCTTAAAGAGAAATGTCAATCGTCAATG 9188
QY 762 ACCTTGCTGCGCAGCTTACCACTTTGGGCACTTCACTCTGAGTCACTTCACTGAGCTC 821
Db 9189 ACCTTGCAATCAGCTGACCACTGGGCACTTCACTCTGAGTCACTTCACTTAACTCAGACT 9248
QY 822 TGAAGACCTTGAACACCAAGATGGAAGCTTCTGAGGCTGGCGCTGAGGAGCCAGTCAAGC 881
Db 9249 TGAAGATCTGAATACCAAGATGGAAGCTTCTGAGTGGCTGTGAGGAGCCGCTGAGAC 9308
QY 882 AGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTGAGCACTTTCTTTTCCAGCTCTG 941
Db 9309 AGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTGAGCACTTTCTTTTCCAGCTCTG 9368
QY 942 TCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCAAG 1001
Db 9369 TTCAAGGCTCCCTGGGAGAGAGCCATCTCACCAACAAAGTGCCTACTATATCAACCAAG 9428
QY 1002 AGACTCAAACTTGTGTGGAGCCATCCCAAAATGACAGAGCTTACCAAGCTTTTAGCTG 1061
Db 9429 AGACCCAAACAACTTGTGTGGAGCCATCCCAAAATGACAGAGCTTACCAAGCTTTTAGCTG 9488
QY 1062 ACCTGAATATGTCAGATTTCTGAGCTTATAGGACTGCCATGAATCTCCAGAGACTCGACA 1121

Qy 1482 TTCTGCATGATTCTATCCAA 1501
|||||
Db 9909 TTCTGCATGATTCTATCCAA 9928
|||||

Search completed: September 19, 2004, 08:30:04
Job time : 3977.06 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2004, 07:39:06 ; Search time 408.027 Seconds
(without alignments)
15627.746 Million cell updates/sec

Title: US-09-845-416-12_COPY_1500_3000

Perfect score: 1501

Sequence: 1 agaagatctagaacaagaac.....ttctgcatttatcccaa 1501

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1501	100.0	3510	6 AAD37240	Human dys
2	1501	100.0	4476	6 AAD37259	Adeno-ass
3	1328.8	88.5	2169	6 AAD37232	Human dys
4	1328.8	88.5	3531	6 AAD37238	Human dys
5	1328.8	88.5	3858	6 AAD37237	Human dys
6	1328.8	88.5	3999	6 AAD37234	Human dys
7	1328.8	88.5	4182	6 AAD37230	Human dys
8	1328.8	88.5	4498	6 AAD37258	Adeno-ass
9	1328.8	88.5	4825	6 AAD37257	Adeno-ass
10	1328.8	88.5	4848	6 AAD37263	Adeno-ass
11	1328.8	88.5	4966	6 AAD37256	Adeno-ass
12	1328.8	88.5	4990	6 AAD37262	Adeno-ass
13	1328.8	88.5	5060	6 AAD37264	Adeno-ass
14	1328.8	88.5	5149	6 AAD37255	Adeno-ass
15	1328.8	88.5	5339	6 ABK81998	DNA encod
16	1328.8	88.5	5462	6 ABK81999	DNA encod
17	1328.8	88.5	5952	5 AAD05794	Human dys
18	1328.8	88.5	8689	6 ABK82000	DNA encod
19	1328.8	88.5	11058	6 AAD37229	Human dys
20	1328.8	88.5	11241	6 ABK82005	CDNA encod
21	1328.8	88.5	11443	6 ABK82002	DNA encod
22	1328.8	88.5	12923	1 AAN90338	Sequence
23	1328.8	88.5	13957	6 ABK81959	CDNA encod

24	1328.8	88.5	13957	6 ABT10904	Human bre
25	1328.8	88.5	13957	6 ABN95786	Gene #228
26	1328.8	88.5	13957	6 ABS69900	Human dys
27	1324	88.2	1821	6 AAD37241	Human dys
28	1317.8	87.8	13977	6 ABS70403	Human bon
29	1149.6	76.6	13815	6 ABK81960	CDNA encod
30	1149.6	76.6	13815	6 ABT199799	Mouse isc
31	1149.6	76.6	13907	6 AAT27558	Shuttle v
32	1148	76.5	13815	2 AAV18885	Mus muscu
33	1041	69.4	3446	6 AAD37242	Human dys
34	1041	69.4	4414	6 AAD37260	Adeno-ass
35	1011.6	67.4	5417	6 ABK81997	DNA encod
36	1006.2	67.0	4402	3 AAZ48567	A rod sho
37	937	62.4	1434	6 AAD37243	Human dys
38	807	53.8	4402	3 AAZ48568	A rod sho
39	671.4	44.7	4075	3 AAZ48569	A rod sho
40	667.2	44.5	3747	3 AAZ48566	A rod sho
41	613.6	40.9	3163	3 AAZ48571	A rod sho
42	555.6	37.0	10705	7 ABT41896	Toxicity
43	554.2	36.9	11096	6 ABK81962	CDNA encod
44	547.4	36.5	6045	2 AAT74665	Utrrophin
45	547.4	36.5	6059	4 AAF84673	Nucleotid

ALIGNMENTS

RESULT 1

AAD37240
ID AAD37240 standard; DNA; 3510 BP.

XX
AC AAD37240;

XX
DT 21-AUG-2002 (first entry)

XX
DE Human dystrophin minigene delta3510.

XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

XX
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX
KW Becker muscular dystrophy; ds.

XX
OS Homo sapiens.

XX
PN WO200183695-A2.

XX
PD 08-NOV-2001.

XX
PF 27-APR-2001; 2001WO-US013677.

XX
PR 28-APR-2000; 2000US-0200777P.

XX
XX (XIAO/) XIAO X.

XX
PI Xiao X;

XX
DR WPI; 2002-049342/06.

XX
PT New dystrophin minigene for treating Duchenne or Becker muscular

XX
PT dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene.

XX
PS Example 1; Page 51-52; 71pp; English.

XX
CC The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3510 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1 and R2), 8407-10227 (rods R23 and R24, hinge H4 and
CC CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ Sequence 3510 BP; 1073 A; 787 C; 828 G; 822 T; 0 U; 0 Other;
Query Match 100.0%; Score 1501; DB 6; Length 3510;
Best Local Similarity 100.0%; Pred No 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAAGATCTAGAACAAAGTCAAGGTCAGGTCATCTCTCACTCACAATGCTGTGTAGT 60
Db 1500 AGAAGATCTAGAACAAAGTCAAGGTCAGGTCATCTCTCACTCACAATGCTGTGTAGT 1559
QY 61 TGATGAATCTAGTGGAGATCAGCACTGCTGCTTTGGAGAACAACTTAAGGTATGGG 120
Db 1560 TGATGAATCTAGTGGAGATCAGCACTGCTGCTTTGGAGAACAACTTAAGGTATGGG 1619
QY 121 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAG 180
Db 1620 AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAG 1679
QY 181 TTCTGACAGTGGAGGCTGTGACCTTTCTCTGAGGAACTTCTGGTGGCTACAGCT 240
Db 1680 TTCTGACAGTGGAGGCTGTGACCTTTCTCTGAGGAACTTCTGGTGGCTACAGCT 1739
QY 241 GAAAGATGATGAATTAAGCGCGGAGGACCTATTGGAGGCGACTTTCCAGCAGTTTCAAGAA 300
Db 1740 GAAAGATGATGAATTAAGCGCGGAGGACCTATTGGAGGCGACTTTCCAGCAGTTTCAAGAA 1799
QY 301 GCAGAACATGATCATAGGCGCTTCAAGAGGGAATTAAGAACTTAAGAACTTGAATCAT 360
Db 1800 GCAGAACATGATCATAGGCGCTTCAAGAGGGAATTAAGAACTTAAGAACTTGAATCAT 1859
QY 361 GAGTACTCTTGAGACTGTAGCAATATTTCTGACAGAGCAGCCCTTTGAAAGGACTAGAGAA 420
Db 1860 GAGTACTCTTGAGACTGTAGCAATATTTCTGACAGAGCAGCCCTTTGAAAGGACTAGAGAA 1919
QY 421 ACTTACAGAGAGCCAGAGAGTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCT 480
Db 1920 ACTTACAGAGAGCCAGAGAGTGCCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCT 1979
QY 481 TCTACGAAAGCAGCTCAGAGGTCATATCTAGTGGGAAATTAAGCTGCACCTCCGC 540
Db 1980 TCTACGAAAGCAGCTCAGAGGTCATATCTAGTGGGAAATTAAGCTGCACCTCCGC 2039
QY 541 TGACTGCGAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCAC 600
Db 2040 TGACTGCGAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCAC 2099
QY 601 GGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTATCAAGGATCTCTGGCAGCCGT 660
Db 2100 GGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTATCAAGGATCTCTGGCAGCCGT 2159
QY 661 GGGGGAATCTCTCACTCTCTCAAGATCACTCTCGAGAAAGTCAAGGCACTTCCAGG 720
Db 2160 GGGGGAATCTCTCACTCTCTCAAGATCACTCTCGAGAAAGTCAAGGCACTTCCAGG 2219
QY 721 AGAAATTCGCGCTCTGAAAGAGAACGTTGAGCCACGTCATGACCTTGTCTGCGCAGTTAC 780
Db 2220 AGAAATTCGCGCTCTGAAAGAGAACGTTGAGCCACGTCATGACCTTGTCTGCGCAGTTAC 2279
QY 781 CACTTTGGGATTCAGCTCTCAGGTATACCTTACCTCAGCCTCTGGAAGACCTGACACACAG 840
Db 2280 CACTTTGGGATTCAGCTCTCAGGTATACCTTACCTCAGCCTCTGGAAGACCTGACACACAG 2339
QY 841 ATGGAAGCTTTCTGAGGTGGCCCTCGAGGACCGAGTCAGGACCTGCATGAAGCCACAG 900
Db 2340 ATGGAAGCTTTCTGAGGTGGCCCTCGAGGACCGAGTCAGGACCTGCATGAAGCCACAG 2399
QY 901 GGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTGTCAGGTCCTCTGGAGAG 960

Db 2400 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCTCAGGGTCCCTGGAGAG 2459
QY 961 AGCCATCTCGCCAAACAAGTGCCTACTATATCAACACAGAGACTCAAAACAATTTGCTG 1020
Db 2460 AGCCATCTCGCCAAACAAGTGCCTACTATATCAACACAGAGACTCAAAACAATTTGCTG 2519
QY 1021 GGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATATGTGAGATT 1080
Db 2520 GGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATATGTGAGATT 2579
QY 1081 CTCAGCTTATAGGACTGCGCATGAAATCCGAAAGCTGCGAAGGCCCTTTGCTTGGATCT 1140
Db 2580 CTCAGCTTATAGGACTGCGCATGAAATCCGAAAGCTGCGAAGGCCCTTTGCTTGGATCT 2639
QY 1141 CTTGAGCCTGTGACGTGATGCTTGGACCGACACAACTCAAGCAAAATGACCA 1200
Db 2640 CTTGAGCCTGTGACGTGATGCTTGGACCGACACAACTCAAGCAAAATGACCA 2699
QY 1201 GCCCATGGATATCTCGAGATTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCA 1260
Db 2700 GCCCATGGATATCTCGAGATTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCA 2759
QY 1261 AGAGCAGCAAAATTTGGTCAAGCTCCCTCTCTGGTGGATATGTCTGACTGGCTGCT 1320
Db 2760 AGAGCAGCAAAATTTGGTCAAGCTCCCTCTCTGGTGGATATGTCTGACTGGCTGCT 2819
QY 1321 GAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGTCTCTGTTCTTTTAAACCTGGCAT 1380
Db 2820 GAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGTCTCTGTTCTTTTAAACCTGGCAT 2879
QY 1381 CATTTCCCTGTGTAAGACATTTGGAAAGACAACTGAGATACATCTTTCAAGCAAGTGGC 1440
Db 2880 CATTTCCCTGTGTAAGACATTTGGAAAGACAACTGAGATACATCTTTCAAGCAAGTGGC 2939
QY 1441 AAGTTCAACAGGATTTGTGACGAGCGAGGCTGGGCTCTCTCTGATGATTTATATCCA 1500
Db 2940 AAGTTCAACAGGATTTGTGACGAGCGAGGCTGGGCTCTCTCTGATGATTTATATCCA 2999
QY 1501 A 1501
Db 3000 A 3000
RESULT 2
AAD37259
ID AAD37259 standard; DNA; 4476 BP.
XX
AC AAD37259;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-3510.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX

DR WPI; 2002-049342/06.
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
XX gene.
XX
PS Example 1; Page 63-65; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ Sequence 4476 BP; 1252 A; 1096 C; 1127 G; 1001 T; 0 U; 0 Other;
Query Match 100.0%; Score 1501; DB 6; Length 4476;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAAGATCTAGAACAAAGTCAGGTCATCTCTCTCACTCACATGTTGGTGTAGT 60
DB 2256 AGAAGATCTAGAACAAAGTCAGGTCATCTCTCTCACTCACATGTTGGTGTAGT 2315
QY 61 TGATGAATCTAGTGGAGATCAGCAACTGTGCTTTGGAAAGAACAACTTAAGGTATTGGG 120
DB 2316 TGATGAATCTAGTGGAGATCAGCAACTGTGCTTTGGAAAGAACAACTTAAGGTATTGGG 2375
QY 121 AGATCGATGGCAACATCTGTAGATGGACAGAGACCGTGGTCTTTTCAAGACAG 180
DB 2376 AGATCGATGGCAACATCTGTAGATGGACAGAGACCGTGGTCTTTTCAAGACAG 2435
QY 181 TTCTGACGAGTGAAGCGTCTGCACCTTTCTCTCAGGAACTTCTGTTGGGTACAGCT 240
DB 2436 TTCTGACGAGTGAAGCGTCTGCACCTTTCTCTCAGGAACTTCTGTTGGGTACAGCT 2495
QY 241 GAAAGATGATGAATTAAGCGGCGAGGACCTATTGGAGGGGACCTTCCAGCAGTTGAGAA 300
DB 2496 GAAAGATGATGAATTAAGCGGCGAGGACCTATTGGAGGGGACCTTCCAGCAGTTGAGAA 2555
QY 301 GCAGAACGATGATACATAGGGCTTCAAGAGGGAATTGAAACAACTGTAATCAT 360
DB 2556 GCAGAACGATGATACATAGGGCTTCAAGAGGGAATTGAAACAACTGTAATCAT 2615
QY 361 GAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTGGAAGGACTTAGAGAA 420
DB 2616 GAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTGGAAGGACTTAGAGAA 2675
QY 421 ACTCTACAGAGCGCCAGAGCTGCGCTCTGAGAGAGAGCCAGAAATGTCACCTGGCT 480
DB 2676 ACTCTACAGAGCGCCAGAGCTGCGCTCTGAGAGAGAGCCAGAAATGTCACCTGGCT 2735
QY 481 TCTACGAAGCAGGCTGAGAGGTCATATCTAGTGGGAAAATTGAACCTGCACTCCGC 540
DB 2736 TCTACGAAGCAGGCTGAGAGGTCATATCTAGTGGGAAAATTGAACCTGCACTCCGC 2795
QY 541 TGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCAC 600
DB 2796 TGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCAC 2855
QY 601 GGATGAGCTGACCTCAAGCTGCCCAAGCTGAGGTGATCAAGGATCCCTGGCAGGCCGT 660
DB 2856 GGATGAGCTGACCTCAAGCTGCCCAAGCTGAGGTGATCAAGGATCCCTGGCAGGCCGT 2915
QY 661 GGGCGATCTCTCATATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG 720

DB 2916 GGGCGATCTCTCATTTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG 2975
QY 721 AGAATTTGGGCTCTGAAAGAGAGCTGAGCGCAGTCAATGACCTTGTCTGCCAGCTTAC 780
DB 2976 AGAATTTGGGCTCTGAAAGAGAGAGCTGAGCGCAGTCAATGACCTTGTCTGCCAGCTTAC 3035
QY 781 CACTTTTGGGCACTTCACTCTCACTCTCACTCTCACTCTGGAAGACCTTGAACACCAG 840
DB 3036 CACTTTTGGGCACTTCACTCTCACTCTCACTCTGGAAGACCTTGAACACCAG 3095
QY 841 ATGGAAGCTTTGCAAGTGGCGCTGAGGAGCGAGTCAAGGAGCTGATGAAGCCCAAG 900
DB 3096 ATGGAAGCTTTGCAAGTGGCGCTGAGGAGCGAGTCAAGGAGCTGATGAAGCCCAAG 3155
QY 901 GGACTTTGGTCCAGCATCTCAGCAGCTTCTTTCACAGCTCTGTCAGGGTCCCTTGGGAGAG 960
DB 3156 GGACTTTGGTCCAGCATCTCAGCAGCTTCTTTCACAGCTCTGTCAGGGTCCCTTGGGAGAG 3215
QY 961 AGCCATCTCGCCCAAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACAACTTGTCTG 1020
DB 3216 AGCCATCTCGCCCAAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACAACTTGTCTG 3275
QY 1021 GGACCATCCCAAAATGACAGAGCTCTACAGCTCTTGTGCTGACCTGATATGTCAGATT 1080
DB 3276 GGACCATCCCAAAATGACAGAGCTCTACAGCTCTTGTGCTGACCTGATATGTCAGATT 3335
QY 1081 CTCAGCTTATAGGACTGCGCATGAAACTCCGAAGACTGCAAGAGCCCTTTTGTGCTGATCT 1140
DB 3336 CTCAGCTTATAGGACTGCGCATGAAACTCCGAAGACTGCAAGAGCCCTTTTGTGCTGATCT 3395
QY 1141 CTTGAGCCTGTGAGCTGATGATGCTTGGACGAGCAACACCTCAAGCAAAATGACCA 1200
DB 3396 CTTGAGCCTGTGAGCTGATGATGCTTGGACGAGCAACACCTCAAGCAAAATGACCA 3455
QY 1201 GCCCATGATATCTCGAGATTATTAATTGTTGACCACTATTTATGACCGCTCGAGCA 1260
DB 3456 GCCCATGATATCTCGAGATTATTAATTGTTGACCACTATTTATGACCGCTCGAGCA 3515
QY 1261 AGAGCACAAATTTGGTCAAGCTCCCTCTCTGCTGGATATGCTCTGAACTGGCTGCT 1320
DB 3516 AGAGCACAAATTTGGTCAAGCTCCCTCTCTGCTGGATATGCTCTGAACTGGCTGCT 3575
QY 1321 GAATGTTTATGATAGCGGACGAGGAGGATCCGTCTGCTCTTTTAAACCTGGCAT 1380
DB 3576 GAATGTTTATGATAGCGGACGAGGAGGATCCGTCTGCTCTTTTAAACCTGGCAT 3635
QY 1381 CATTTCCTCTGTATAAGCAGCATTTTGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC 1440
DB 3636 CATTTCCTCTGTATAAGCAGCATTTTGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC 3695
QY 1441 AAGTTCAACAGGATTTTGTGACCGAGGAGCTGGGCTCTCTCTGCAATGATTTCTATCCA 1500
DB 3696 AAGTTCAACAGGATTTTGTGACCGAGGAGCTGGGCTCTCTCTGCAATGATTTCTATCCA 3755
QY 1501 A 1501
DB 3756 A 3756

RESULT 3

AAD37232

ID AAD37232 standard; DNA; 2169 BP.

XX

AC AAD37232;

XX

DT 21-AUG-2002 (first entry)

XX

Human dystrophin rod, hinge and CR domain regions encoding DNA #1.

XX

Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

XX

adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX

Becker muscular dystrophy; ds.

KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
OS Homo sapiens.
XX
XX WO200183695-A2.
XX
XX
XX PD 08-NOV-2001.
XX
XX 27-APR-2001; 2001WO-US013677.
XX
XX
XX PR 28-APR-2000; 2000US-020077P.
XX
XX (XIAO/) XIAO X.
XX
XX Xiao X;
XX
XX WPI; 2002-049342/06.
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
XX Example 1; Page 50-51; 71pp; English.
XX
XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus,
CC hinge H1 and rod R1), 8059-10227 (rods R22, R23 and R24, hinge H4 and CR
CC domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 U; 0 Other;

Query Match 88.5%; Score 1328.8; DB 6; Length 3531;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 162 GGGTTCTTTTACAAAGACAGTTCTGACCAAGTGGAGGGTCTGACCTTTCTCTCCAGGAAC 221
DB 1682 GGTCCCATTTTGAAGCCAGTTCTGACCAAGTGGAGGGTCTGACCTTTCTCTCCAGGAAC 1741

QY 222 TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGG 281
DB 1742 TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGG 1801

QY 282 ACTTTCAGCAGTTCAGAAAGCAGAACGATGATCATAGGGCTTCAAGAGGAAATTGAAA 341
DB 1802 ACTTTCAGCAGTTCAGAAAGCAGAACGATGATCATAGGGCTTCAAGAGGAAATTGAAA 1861

QY 342 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGC 401
DB 1862 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGC 1921

QY 402 CTTTGGAGGACTAGAGAAACTTACCCAGGAGCCCGAGAGAGTCCCTCTCTGAGGAGAGAG 461
DB 1922 CTTTGGAGGACTAGAGAAACTTACCCAGGAGCCCGAGAGAGTCCCTCTCTGAGGAGAGAG 1981

QY 462 CCCAGAAATGTCACCTCGCGTTCTACGAAAGCAGGCTGAGGAGGTCAATACATGAGTGGGAAA 521
DB 1982 CCCAGAAATGTCACCTCGCGTTCTACGAAAGCAGGCTGAGGAGGTCAATACATGAGTGGGAAA 2041

QY 522 AATTGAACTGTCACCTCGCGTTCTGAGTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC 581

RESULT 5
AAD37237
ID AAD37237 standard; DNA; 3858 BP.
XX
XX AAD37237;
XX

DB 2042 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC 2101
QY 582 AGGAACCTTTCAAGAGGCCACCGATGAGCTTCAAGCTGCGCAAGCTGAGGTGATCA 641
DB 2102 AGGAACCTTTCAAGAGGCCACCGATGAGCTTCAAGCTGCGCAAGCTGAGGTGATCA 2161
QY 642 AGGATCTCTGCGACGCCGTGGCGCATCTCTCATTTGACTCTCTCAAGATCACTCTCGAGA 701
DB 2162 AGGATCTCTGCGACGCCGTGGCGCATCTCTCATTTGACTCTCTCAAGATCACTCTCGAGA 2221
QY 702 AAGTCAAGGCACTTCGAGGAGAAAATTGGCCCTCTGAAAAGAGAACTGAGCCACCTCAATG 761
DB 2222 AAGTCAAGGCACTTCGAGGAGAAAATTGGCCCTCTGAAAAGAGAACTGAGCCACCTCAATG 2281
QY 762 ACCTTGCTCGCCAGCTTACCACTTTGGCATTTAGCTCTTACCGTATTAACCTGACACTC 821
DB 2282 ACCTTGCTCGCCAGCTTACCACTTTGGCATTTAGCTCTTACCGTATTAACCTGACACTC 2341
QY 822 TGGAAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCCGTGAGACCCGAGTCAAGGC 881
DB 2342 TGGAAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCCGTGAGACCCGAGTCAAGGC 2401
QY 882 AGCTGCATGAAGCCCAAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGGTCTG 941
DB 2402 AGCTGCATGAAGCCCAAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGGTCTG 2461
QY 942 TCCAGGCTCCCTGGGAGAGCCCATCTCGCCAAACAAAGTCCCTACTATATCAACACG 1001
DB 2462 TCCAGGCTCCCTGGGAGAGCCCATCTCGCCAAACAAAGTCCCTACTATATCAACACG 2521
QY 1002 AGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTTACAGTCTTTAGCTG 1061
DB 2522 AGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTTACAGTCTTTAGCTG 2581
QY 1062 ACCTGAATTAATGTGAGTCTCAGCTTATAGAGTCCATGAACTCCGAGAGCTGCAGA 1121
DB 2582 ACCTGAATTAATGTGAGTCTCAGCTTATAGAGTCCATGAACTCCGAGAGCTGCAGA 2641
QY 1122 AGGCCCTTTGCTTGGATCTCTTTGAGCCCTGTCAGCTGTCATGTCCTTTGACAGCAGACA 1181
DB 2642 AGGCCCTTTGCTTGGATCTCTTTGAGCCCTGTCAGCTGTCATGTCCTTTGACAGCAGACA 2701

QY 1182 ACCTCAAGCAAAATGACCCAGCCATGGATATCTCTGAGATTATTAATTTGTTGACCCTA 1241
DB 2702 ACCTCAAGCAAAATGACCCAGCCATGGATATCTCTGAGATTATTAATTTGTTGACCCTA 2761
QY 1242 TTTATGACCGCTGGAGCAAGAGCACAACAATTTTGTCAAGTCCCTCTCTGCGTGGATA 1301
DB 2762 TTTATGACCGCTGGAGCAAGAGCACAACAATTTTGTCAAGTCCCTCTCTGCGTGGATA 2821
QY 1302 TGTGCTGAACCTGGCTGCTGTAATGTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 1361
DB 2822 TGTGCTGAACCTGGCTGCTGTAATGTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 2881
QY 1362 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGGACATTTGGAAGACAAAGTACAGAT 1421
DB 2882 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGGACATTTGGAAGACAAAGTACAGAT 2941
QY 1422 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGAGGTGGGCTCC 1481
DB 2942 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGAGGTGGGCTCC 3001
QY 1482 TTCTGCATGATTCTATCCAA 1501
DB 3002 TTCTGCATGATTCTATCCAA 3021

DT 21-AUG-2002 (first entry)
XX Human dystrophin minigene delta3849.
XX
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
FI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
XX dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
XX gene.
PS Example 1; Page 48-49; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 U; 0 Other;

Query Match 88.5%; Score 1328.8; DB 6; Length 3858;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 162 GGGTTCTTTACAGACAGTTCTGACCAGTGGAGCGTCTGCACCTTTCTCTGCAGGAAC 221
DB 2009 GGTCCTCAATTTGGAGCGAGTTCTGACAGTGGAGCGTCTGCACCTTTCTCTGCAGGAAC 2068

QY 222 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG 281
DB 2069 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG 2128

QY 282 ACTTTCCAGCAGTTCAGACGACGATGATACATAGGCGCTTCAGAGGGGAATTAAGA 341
DB 2129 ACTTTCCAGCAGTTCAGACGACGATGATACATAGGCGCTTCAGAGGGGAATTAAGA 2188

QY 342 CTAAGAAGACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTCTGCAGAGCAGC 401
DB 2189 CTAAGAAGACCTGTAATCATGAGTACTCTTTGAGACTGTACGAATATTCTGCAGAGCAGC 2248

QY 402 CTTTGGAGAGCTAGAGAACTCTACAGAGCCCGAGAGCTGCGCTCTCTGAGGAGAGAG 461
DB 2249 CTTTGGAGAGCTAGAGAACTCTACAGAGCCCGAGAGCTGCGCTCTCTGAGGAGAGAG 2308

QY 462 CCAGAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATACCTGAGTGGGAAA 521
AAG37234

DB 2309 CCCAGAAATGTCACTCGGCTTCTACGAAAGCAGCGTGGAGGTTCAATACTGAGTGGGAAA 2368
QY 522 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCC 581
DB 2369 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCC 2428
QY 582 AGAACTTTCAAGAGAGCCACGATGAGCTGAGACCTCAAGCTGCGCAAGCTGAGGTGATCA 641
DB 2429 AGAACTTTCAAGAGAGCCACGATGAGCTGAGACCTCAAGCTGCGCAAGCTGAGGTGATCA 2488
QY 642 AGGATCTCTGCGAGCCCGTGGCGATCTCTCATTTGATCTCTCCAGATCACTCCGAGA 701
DB 2489 AGGATCTCTGCGAGCCCGTGGCGATCTCTCATTTGATCTCTCCAGATCACTCCGAGA 2548
QY 702 AAGTCAAGGCACTTCGAGGAGAAAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCGAATG 761
DB 2549 AAGTCAAGGCACTTCGAGGAGAAAAATTGGCCCTCTGAAAGAGAACGTGAGCCACGTCGAATG 2608
QY 762 ACCTTGCTGCGCAGCTTACCACTTTGGCATTGAGCTCTCACCCTATTAACCTCAGCACTC 821
DB 2609 ACCTTGCTGCGCAGCTTACCACTTTGGCATTGAGCTCTCACCCTATTAACCTCAGCACTC 2668
QY 822 TGGAAAGCTTCAACACAGATGGAAAGCTTCTGAGGTGGCGCTCGAGACCGAGTCAGGC 881
DB 2669 TGGAAAGCTTCAACACAGATGGAAAGCTTCTGAGGTGGCGCTCGAGACCGAGTCAGGC 2728
QY 882 AGCTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCATCTTTTCCACGCTG 941
DB 2729 AGCTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCATCTTTTCCACGCTG 2788
QY 942 TCCAGGCTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCAACG 1001
DB 2789 TCCAGGCTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCAACG 2848
QY 1002 AGACTCAAAACAACTTGTGGACCACTCCCAAAATGACAGAGCTTACCAGCTTTAGCTG 1061
DB 2849 AGACTCAAAACAACTTGTGGACCACTCCCAAAATGACAGAGCTTACCAGCTTTAGCTG 2908
QY 1062 ACCTGATGAATGTGAGATCTCTCAGCTTATAGGCTGCCATGAACCTCCGAGAGCTGAGA 1121
DB 2909 ACCTGATGAATGTGAGATCTCTCAGCTTATAGGCTGCCATGAACCTCCGAGAGCTGAGA 2968
QY 1122 AGGCCCCCTTGTGGATCTCTTGCAGCTGTGATGCTGATGCTTGGACAGCACA 1181
DB 2969 AGGCCCCCTTGTGGATCTCTTGCAGCTGTGATGCTGATGCTTGGACAGCACA 3028
QY 1182 ACCTCAAGCAAAATGACAGCCCATGATATCTGAGATTTATATTTGTTGACCACTA 1241
DB 3029 ACCTCAAGCAAAATGACAGCCCATGATATCTGAGATTTATATTTGTTGACCACTA 3088
QY 1242 TTTATGACCGCTGGAGCAAGAGCACAACAAATTTGGTCAACGCTCCCTCTCTGGGTGGATA 1301
DB 3089 TTTATGACCGCTGGAGCAAGAGCACAACAAATTTGGTCAACGCTCCCTCTCTGGGTGGATA 3148
QY 1302 TGTGTCTGAATGCTGCTGCTGAAATTTATGATACGGGACGAAAGGAGGATCCGTGTCC 1361
DB 3149 TGTGTCTGAATGCTGCTGCTGAAATTTATGATACGGGACGAAAGGAGGATCCGTGTCC 3208
QY 1362 TGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAGCACAATTTTGGAGAGCAAGTACAGAT 1421
DB 3209 TGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAGCACAATTTTGGAGAGCAAGTACAGAT 3268
QY 1422 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGGCTCC 1481
DB 3269 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGGCTCC 3328
QY 1482 TTCTGCAATGTTCTATCCAA 1501
DB 3329 TTCTGCAATGTTCTATCCAA 3348

RESULT 6
AAD37234

AD37234 standard; DNA; 3999 BP.
AC AAD37234;
XX
XX
DT 21-AUG-2002 (first entry)
XX
XX Human dystrophin minigene delta3990.
XX DE
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; GB.
XX OS
XX Homo sapiens.
XX WO200183695-A2.
XX PN
XX 08-NOV-2001.
XX PD
XX 27-APR-2001; 2001WO-US013677.
XX PF
XX 28-APR-2000; 2000US-0200777P.
XX PR
XX (XIAO/) XIAO X.
XX PA
XX Xiao X;
XX PI
XX WPI; 2002-049342/06.
XX DR
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
XX Example 1; Page 46-47; 71pp; English.
XX
XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
XX Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 U; 0 Other;
XX
XX Query Match 88.5%; Score 1328.8; DB 6; Length 3999;
XX Best Local Similarity 99.5%; Pred. No. 0;
XX Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
XX 162 GGGTCTTTTACAGAGAGTCTCACCAGTGGAGCGCTCTGCACCTTCTCTCGAGGAAC 221
XX
XX 2150 GGTCCCAATTTGAAGCCAGTCTGACAGTGGAGCGCTCTGCACCTTCTCTCGAGGAAC 2209
XX
XX 222 TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAACCGGACGACCTATTGGAGCG 281
XX
XX 2210 TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAACCGGACGACCTATTGGAGCG 2269
XX
XX 282 ACTTTCAGAGGTTTCAGAGCAGACGATGTACATAGGCTTCAAGAGGGAATTTGAAA 341
XX
XX 2270 ACTTTCAGAGGTTTCAGAGCAGACGATGTACATAGGCTTCAAGAGGGAATTTGAAA 2329
XX
XX 342 CTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAG 401
XX
XX 2330 CTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAG 2389
XX
XX 402 CTTTGGAGAGGCTAGAGAACTCTACAGAGGCCACAGAGCTCCCTCTGAGGAGAGAG 461
XX

Db 2390 CTTTGGAGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTCCCTCTGAGAGAGAG 2449
Qy 462 CCCAGAAATGTCACCTCGCTTCTACAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA 521
Db 2450 CCCAGAAATGTCACCTCGCTTCTACAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA 2509
Qy 522 AATTGAACCTGCATCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAAGACTCC 581
Db 2510 AATTGAACCTGCATCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAAGACTCC 2569
Qy 582 AGGAACTTCAGAGGCGCACGATGAGTGGACCTCAAGCTGCGCAAGCTGAGTGTATCA 641
Db 2570 AGGAACTTCAGAGGCGCACGATGAGTGGACCTCAAGCTGCGCAAGCTGAGTGTATCA 2629
Qy 642 AGGATCTCTGACGCGCTGGCGCATCTCTCATTTGATCTCTCTCAAGATCACTCTGAGA 701
Db 2630 AGGATCTCTGACGCGCTGGCGCATCTCTCATTTGATCTCTCTCAAGATCACTCTGAGA 2689
Qy 762 AAGTCAAGGCACTTCGAGGAGAAATTCGCTCTGAAAGAGAACTGAGCAGCTCAATG 761
Db 2690 AAGTCAAGGCACTTCGAGGAGAAATTCGCTCTGAAAGAGAACTGAGCAGCTCAATG 2749
Qy 762 ACCTTGTCTCGCAGCTTACCACTTTGGGCATTTGAGCTCTCAACCTTAACTCAGCACTC 821
Db 2750 ACCTTGTCTCGCAGCTTACCACTTTGGGCATTTGAGCTCTCAACCTTAACTCAGCACTC 2809
Qy 822 TGGAAAGCCTGAACACACAGATGGAAGCTTCGAGGTGGCCCTGAGGACCGAGTCAGGC 881
Db 2810 TGGAAAGCCTGAACACACAGATGGAAGCTTCGAGGTGGCCCTGAGGACCGAGTCAGGC 2869
Qy 882 AGCTGCATGAAGCCACAGGACCTTTGGTCCAGATCTCAGACATCTCTTCCACGCTG 941
Db 2870 AGCTGCATGAAGCCACAGGACCTTTGGTCCAGATCTCAGACATCTCTTCCACGCTG 2929
Qy 942 TCCAGGCTCCCTGGAGAGGCCATCTCGCAAAACAAAGTGCCTTATATCAACCAAG 1001
Db 2930 TCCAGGCTCCCTGGAGAGGCCATCTCGCAAAACAAAGTGCCTTATATCAACCAAG 2989
Qy 1002 AGACTCAACAACTTGTCTGGGACCATCCAAATAGACAGAGCTCTACAGTCTTTAGCTG 1061
Db 2990 AGACTCAACAACTTGTCTGGGACCATCCAAATAGACAGAGCTCTACAGTCTTTAGCTG 3049
Qy 1062 ACCTGAATAATGTCAAGATCTCAGCTTATAGGAGTGCATGAACTCCGAAGACTGCAGA 1121
Db 3050 ACCTGAATAATGTCAAGATCTCAGCTTATAGGAGTGCATGAACTCCGAAGACTGCAGA 3109
Qy 1122 AGGCTCTTGTCTTGTGATCTCTTGAGCTCTGAGCTGATGTGATGCTTGGACAGCACA 1181
Db 3110 AGGCTCTTGTCTTGTGATCTCTTGAGCTCTGAGCTGATGTGATGCTTGGACAGCACA 3169
Qy 1182 ACCTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTAATTAATTTGACCACTA 1241
Db 3170 ACCTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTAATTAATTTGACCACTA 3229
Qy 1242 TTTATGACCGCTCGAGCAAGACACAAATTTGGTCAAGCTCCCTCTCTGCGTGGATA 1301
Db 3230 TTTATGACCGCTCGAGCAAGACACAAATTTGGTCAAGCTCCCTCTCTGCGTGGATA 3289
Qy 1302 TGTGTCTGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1361
Db 3290 TGTGTCTGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3349
Qy 1362 TGTCTTTTAAAACTGGCATCAATTTCCCTGTGTAAAGCAATTTTGAAGACAAAGTACAGAT 1421
Db 3350 TGTCTTTTAAAACTGGCATCAATTTCCCTGTGTAAAGCAATTTTGAAGACAAAGTACAGAT 3409
Qy 1422 ACCTTTTCAAGCAAGTGGCAAGTTCACAGATTTTGTGACAGCGCAGGCTGGGCTCC 1481
Db 3410 ACCTTTTCAAGCAAGTGGCAAGTTCACAGATTTTGTGACAGCGCAGGCTGGGCTCC 3469
Qy 1482 TTCTGCATCATTTCTATCCAA 1501
Db 3470 TTCTGCATCATTTCTATCCAA 3489

RESULT 7
 AAD37230
 ID AAD37230 standard; DNA; 4182 BP.
 XX
 AC AAD37230;
 XX
 DT 21-AUG-2002 (first entry)
 XX
 DE Human dystrophin minigene delta4173.
 XX
 KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
 KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
 KW Becker muscular dystrophy; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200183695-A2.
 XX
 PD 08-NOV-2001.
 XX
 PF 27-APR-2001; 2001WO-US013677.
 XX
 PR 28-APR-2000; 2000US-0200777P.
 XX
 PA (XIAO/) XIAO X.
 XX
 PI Xiao X;
 XX
 DR WPI; 2002-049342/06.
 XX
 PT New dystrophin minigene for treating Duchenne or Becker muscular
 PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
 PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
 PT gene.
 XX
 PS Example 1; Page 43-44; 71pp; English.
 XX
 CC The present invention relates to an isolated nucleotide sequence encoding
 CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
 CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
 CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
 CC invention also relates to a recombinant adeno-associated virus (AAV)
 CC comprising dystrophin minigene operably linked to an expression control
 CC element. The dystrophin minigene in operable linkage with an expression
 CC control element, in a recombinant adeno-associated virus or retrovirus is
 CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
 CC dystrophy (BMD) in a mammalian subject. The present sequence is human
 CC dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,
 CC hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R2, R3 and R4,
 CC hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
 XX
 SQ Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 U; 0 Other;
 XX
 Query Match 88.5%; Score 1328.8; DB 6; Length 4182;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 162 GGGTTCTTTTACAGACAGTCTGACCAAGTGGAGCGTCTGACACCTTTCTCTGCAGGAAC 221
 DB 2333 GGTCCCATTTGGAAGCAGTCTGACCAAGTGGAGCGTCTGACACCTTTCTCTGCAGGAAC 2392
 QY 222 TTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGAGGCG 281
 DB 2393 TTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGAGGCG 2452
 QY 282 ACTTTCCAGCAGTTCAAGACGAGACCATGTACATAGGCGCTTCAAGAGGGAATTAAGAA 341
 DB 2453 ACTTTCCAGCAGTTCAAGACGAGACCATGTACATAGGCGCTTCAAGAGGGAATTAAGAA 2512
 QY 342 CTAAAGAACTGTAAATCATGACTCTTGTGACATGTACGAATATTCTTGACAGAGCAGC 401

Db 2513 CTAAAGAACTGTAAATCATGACTCTTGTGAGACTGTAGCAATATTCTTGACAGAGCAGC 2572
 QY 402 CTTTGGAAAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCCTCTCTGAGGAGAGAG 461
 Db 2573 CTTTGGAAAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCCTCTCTGAGGAGAGAG 2632
 QY 462 CCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 521
 Db 2633 CCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 2692
 QY 522 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAGACTCC 581
 Db 2693 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGAGACTCC 2752
 QY 582 AGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGTGAATCA 641
 Db 2753 AGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGTGAATCA 2812
 QY 642 AGGATCCCTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCAAGATCACCTCGAGA 701
 Db 2813 AGGATCCCTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCAAGATCACCTCGAGA 2872
 QY 702 AAGTCAAGCACTTCGAGGAGAAAAATTCGCTCTGAAAGAGAACTGAGCCACCTCAATG 761
 Db 2873 AAGTCAAGCACTTCGAGGAGAAAAATTCGCTCTGAAAGAGAACTGAGCCACCTCAATG 2932
 QY 762 ACCTTGCTCGCAGCTTACCACTTTGGGCACTCAGCTCTCACCGTATAAACCCTGAGCACTC 821
 Db 2933 ACCTTGCTCGCAGCTTACCACTTTGGGCACTCAGCTCTCACCGTATAAACCCTGAGCACTC 2992
 QY 822 TGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAAGC 881
 Db 2993 TGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAAGC 3052
 QY 882 AGCTGCATGAAGCCCAAGGACCTTTGGTCCAGCATCTCAGCACTTCTTTCACAGCTCTG 941
 Db 3053 AGCTGCATGAAGCCCAAGGACCTTTGGTCCAGCATCTCAGCACTTCTTTCACAGCTCTG 3112
 QY 942 TCCAGGGTCCCTGGGAGAGGCCATCTCGCAAAACAAAGTGCCCTACTATATCAACCACG 1001
 Db 3113 TCCAGGGTCCCTGGGAGAGGCCATCTCGCAAAACAAAGTGCCCTACTATATCAACCACG 3172
 QY 1002 AGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTAGCTG 1061
 Db 3173 AGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTAGCTG 3232
 QY 1062 ACCTGAAATATGTCTGAGTCTTCTGAGTCTATAGGACTGCATGAAACTCCGAGGACTGAGA 1121
 Db 3233 ACCTGAAATATGTCTGAGTCTTCTGAGTCTATAGGACTGCATGAAACTCCGAGGACTGAGA 3292
 QY 1122 AGGCCCTTTGCTTGGATCTCTTGTGAGCTGTGAGTGCATGTGATGCTTGGACAGACACA 1181
 Db 3293 AGGCCCTTTGCTTGGATCTCTTGTGAGCTGTGAGTGCATGTGATGCTTGGACAGACACA 3352
 QY 1182 ACCTCAAGCAAAATGACCGCCATGATATCTGAGATTAATTTGTTGACCACTA 1241
 Db 3353 ACCTCAAGCAAAATGACCGCCATGATATCTGAGATTAATTTGTTGACCACTA 3412
 QY 1242 TTTATGACCCCTGGAGCAAGAGACACAATTTGGTCAACCGTCCCTCTCTGCGTGGATA 1301
 Db 3413 TTTATGACCCCTGGAGCAAGAGACACAATTTGGTCAACCGTCCCTCTCTGCGTGGATA 3472
 QY 1302 TGTGCTGAATGGCTGCTGAATTTATGATACGGGACGAAACAGGAGGATCCGTGTCC 1361
 Db 3473 TGTGCTGAATGGCTGCTGAATTTATGATACGGGACGAAACAGGAGGATCCGTGTCC 3532
 QY 1362 TGTCTTTTAAACTGGGCACTCAATTTCCCTGTGTAAGACACATTTGGAAGCAAGTACAGAT 1421
 Db 3533 TGTCTTTTAAACTGGGCACTCAATTTCCCTGTGTAAGACACATTTGGAAGCAAGTACAGAT 3592
 QY 1422 ACCTTTTCAAGCAAGTGGGCAAGTTTCAACAGGATTTTGTGACACGAGCTGGGCTCC 1481
 Db 3593 ACCTTTTCAAGCAAGTGGGCAAGTTTCAACAGGATTTTGTGACACGAGCTGGGCTCC 3652

QY 1482 TTCTGCATGATTCTATCCAA 1501
Db |||||
3653 TTCTGCATGATTCTATCCAA 3672
|||
RESULT 8
AAD37258
ID AAD37258 standard; DNA; 4498 BP.
XX AAD37258;
XX
XX 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-3531.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
XX Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
XX WO200183695-A2.
PN
XX
XX 08-NOV-2001.
PD
XX
XX 27-APR-2001; 2001WO-US013677.
PF
XX
XX 28-APR-2000; 2000US-020077P.
PR
XX
XX (XIAO/) XIAO X.
PA
XX
XX Xiao X;
PI
XX
XX WPI; 2002-049342/06.
DR
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
XX Example 1; Page 62-63; 71pp; English.
PS
XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ Sequence 4498 BP; 1251 A; 1118 C; 1123 G; 1006 T; 0 U; 0 Other;
Query Match 88.5%; Score 1328.8; DB 6; Length 4498;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 162 GGGTCTTTTACAGAGAGTCTGACAGTGGAGCGTCTGACCTTCTCCAGGAAC 221
Db |||||
2439 GGTCCCATTTGGAGCGGAGTCTGACAGTGGAGCGTCTGACCTTCTCCAGGAAC 2498
|||
QY 222 TTCGTGTGTGCTACAGTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCG 281
Db |||||
2499 TTCGTGTGTGCTACAGTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCG 2558
|||
QY 282 ACTTCCAGCAGTTCAGAGAGAGAGAGATGATGATACATAGGCGCTTCAAGAGGGAATTGAAA 341
|||

Db 3639 TGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGACACATTTGGAAACACAGTACAGAT 3698
 QY 1422 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGAGATTTTGTACCAAGCGCAGCTGGGCTCC 1481
 Db 3699 ACCTTTCAAGCAAGTGGCAAGTTCAACAGAGATTTTGTACCAAGCGCAGCTGGGCTCC 3758
 QY 1482 TTCTGCATGATTCATCCAA 1501
 Db 3759 TTCTGCATGATTCATCCAA 3778

RESULT 9
 AAD37257
 ID AAD37257 standard; DNA; 4825 BP.
 XX AC
 AC AAD37257;
 XX DT
 XX 21-AUG-2002 (first entry)
 DE Adeno-associated virus vector plasmid, AAV-MCK-delta3849.
 XX KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
 KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
 KW Becker muscular dystrophy; ds.
 XX OS Homo sapiens.
 OS Unidentified.
 OS Chimeric.
 XX FN WO200183695-A2.
 XX PD 08-NOV-2001.
 XX PF 27-APR-2001; 2001WO-US013677.
 XX PR 28-APR-2000; 2000US-0200777P.
 XX PA (XIAO/) XIAO X.
 XX PI Xiao X;
 XX WPI; 2002-049342/06.
 DR New dystrophin minigene for treating Duchenne or Becker muscular
 PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
 PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
 PT gene.
 XX Example 1; Page 61-62; 71pp; English.

XX CC The present invention relates to an isolated nucleotide sequence encoding
 CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
 CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
 CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
 CC invention also relates to a recombinant adeno-associated virus (AAV)
 CC comprising dystrophin minigene operably linked to an expression control
 CC element. The dystrophin minigene in operable linkage with an expression
 CC control element, in a recombinant adeno-associated virus or retrovirus is
 CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
 CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
 CC vector plasmid construct containing human dystrophin minigenes, a muscle
 CC creatine kinase (MCK) promoter and a small polyA signal sequence

XX SQ Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 U; 0 Other;
 Query Match 88.5%; Score 1328.8; DB 6; Length 4825;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 162 GGGTCTTTTACAGACAGTCTGACCAAGTGGAGCGTCTGACCTTTCTCTGCAGGAAC 221
 Db 2766 GGTCCCATTTTGAAGCCAGTTCTGACAGTGGAGCGTCTGACCTTTCTCTGCAGGAAC 2825

QY 222 TTCTGGTGTGCTACAGCTGAAGATGATGATTAAGCCGCGCAGGCACCTATTGGAGCG 281
 Db 2826 TTCTGGTGTGCTACAGCTGAAGATGATGATTAAGCCGCGCAGGCACCTATTGGAGCG 2885
 QY 282 ACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTTGAAAA 341
 Db 2886 ACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTTGAAAA 2945
 QY 342 CTAAGAAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCAGAGCAGC 401
 Db 2946 CTAAGAAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCAGAGCAGC 3005
 QY 402 CTTTGGAGGAGCTAGAGAACTCTACAGAGGCCAGAGAGCTGCCCTCCTGAGGAGAGAG 461
 Db 3006 CTTTGGAGGAGCTAGAGAACTCTACAGAGGCCAGAGAGCTGCCCTCCTGAGGAGAGAG 3065
 QY 462 CCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAAA 521
 Db 3066 CCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAAA 3125
 QY 522 AATTGAACCTGCATCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCC 581
 Db 3126 AATTGAACCTGCATCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCC 3185
 QY 582 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCA 641
 Db 3186 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCA 3245
 QY 642 AGGATCTTGCAGCGCTGGGCGATCTCTCATTTGACTCTCTCCAGATCACCCTCAGA 701
 Db 3246 AGGATCTTGCAGCGCTGGGCGATCTCTCATTTGACTCTCTCCAGATCACCCTCAGA 3305
 QY 702 AAGTCAAGGCACTTTCGAGGAGAAATTCGGCTCTGAAAGAGAACTCAGCCACCTCAATG 761
 Db 3306 AAGTCAAGGCACTTTCGAGGAGAAATTCGGCTCTGAAAGAGAACTCAGCCACCTCAATG 3365
 QY 762 ACCTTGCTGCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAACCCTCAGACTC 821
 Db 3366 ACCTTGCTGCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAACCCTCAGACTC 3425
 QY 822 TGGAGACCTGAACACACAGATGGAAGCTTCTGAGGTTGGCGCTCGAGAGCCGAGTCAAGC 881
 Db 3426 TGGAGACCTGAACACACAGATGGAAGCTTCTGAGGTTGGCGCTCGAGAGCCGAGTCAAGC 3485
 QY 882 AGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTG 941
 Db 3486 AGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTG 3545
 QY 942 TCCAGGTTCCCTGGAGAGAGCCATCTCGCAACAAAGTGCCTACTATATCAACCAAG 1001
 Db 3546 TCCAGGTTCCCTGGAGAGAGCCATCTCGCAACAAAGTGCCTACTATATCAACCAAG 3605
 QY 1002 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGCTCTTTAGCTG 1061
 Db 3606 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGCTCTTTAGCTG 3665
 QY 1062 ACCTGAATATGTCAGATTTCTCAGCTTATAGGACTGCCATGAACTCCGAGAGCTGAGA 1121
 Db 3666 ACCTGAATATGTCAGATTTCTCAGCTTATAGGACTGCCATGAACTCCGAGAGCTGAGA 3725
 QY 1122 AGGCCCTTTGCTGGATCTCTTGGAGCCTGTGAGCTGATGTGCTTGGACCAAGCACA 1181
 Db 3726 AGGCCCTTTGCTGGATCTCTTGGAGCCTGTGAGCTGATGTGCTTGGACCAAGCACA 3785
 QY 1182 ACCTCAAGCAAAATGACAGCCCATGATATCTTCAGATATTAATTTGTTGACCACTA 1241
 Db 3786 ACCTCAAGCAAAATGACAGCCCATGATATCTTCAGATATTAATTTGTTGACCACTA 3845
 QY 1242 TTTATGACCGCTGGAGCAAGAGCACAACAAATTTGGTCAACGCTCCCTCTCTCGGTGGATA 1301
 Db 3846 TTTATGACCGCTGGAGCAAGAGCACAACAAATTTGGTCAACGCTCCCTCTCTCTCGGTGGATA 3905
 QY 1302 TGCTGTGAATGGCTGCTGTAATTTTATGATACGGGACGAAACAGGAGGATCCGTGTCC 1361

```
Db 3906 TGTGTCGTAACGGGTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 3965
Qy 1362 TGTCTTTTAAAACTGGCATCATTTCCCTGTGTGTAAGACACATTTGGAAGACAGTACAGAT 1421
Db 3966 TGTCTTTTAAAACTGGCATCATTTCCCTGTGTGTAAGACACATTTGGAAGACAGTACAGAT 4025
Qy 1422 ACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACGAGCGCAGGCTGGGCGCTCC 1481
Db 4026 ACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACGAGCGCAGGCTGGGCGCTCC 4085
Qy 1482 TTCTGCATGATCTATCCAA 1501
Db 4086 TTCTGCATGATCTATCCAA 4105
```

RESULT 10
AAD37263

ID AAD37263 standard; DNA; 4848 BP.

AC AAD37263;

XX 21-AUG-2002 (first entry)

XX Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.

XX Homo sapiens.

OS Cytomegalovirus.

OS Unidentified.

OS Chimeric.

XX WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US013677.

XX 28-APR-2000; 2000US-0200777P.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/06.

XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.

XX Example 1; Page 68-70; 71pp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomegalovirus (CMV) promoter and a small polyA signal sequence
XX

SQ Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 U; 0 Other;

Query Match

Best Local Similarity 88.5%; Score 1328.8; DB 6; Length 4848;

Pred. No. 0;

```
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 162 GGGTTCTTTTACAGACAGTTCTGACAGTGAAGCGTCTGCACCTTTCTTCAGGAAC 221
Db 2789 GGTCCCATTTTGAAGCGCAGTTCTGACAGTGAAGCGTCTGCACCTTTCTTCAGGAAC 2848
Qy 222 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGGC 281
Db 2849 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGGC 2908
Qy 282 ACTTTCCAGCAGTTTCAAGAGAGAAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAA 341
Db 2909 ACTTTCCAGCAGTTTCAAGAGAGAAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAA 2968
Qy 342 CTAAGAACCTGTAATCATGAGTACTCTTGAAGCTGTACGAATATTTCTGACAGACAGC 401
Db 2969 CTAAGAACCTGTAATCATGAGTACTCTTGAAGCTGTACGAATATTTCTGACAGACAGC 3028
Qy 402 CTTTGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAG 461
Db 3029 CTTTGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAG 3088
Qy 462 CCCAGATGTCACCTGGCTTCTACGAAACAGGCTGAGGAGTCAATATCTGAGTGGGAAA 521
Db 3089 CCCAGATGTCACCTGGCTTCTACGAAACAGGCTGAGGAGTCAATATCTGAGTGGGAAA 3148
Qy 522 AATTGAACCTGCACTCCGCTGACTGCGAGAGAAAATAGATGAGACCTTGTGAAGACTCC 581
Db 3149 AATTGAACCTGCACTCCGCTGACTGCGAGAGAAAATAGATGAGACCTTGTGAAGACTCC 3208
Qy 582 AGGAACCTCAAGAGGCCACGGATGAGTGAACCTCAAGCTGCGCAAGCTGAGGTGATCA 641
Db 3209 AGGAACCTCAAGAGGCCACGGATGAGTGAACCTCAAGCTGCGCAAGCTGAGGTGATCA 3268
Qy 642 AGGGATCTTGGCAGCGCGTGGCGGATCTCTCATTTGACTCTCTCCAAAGATGACCTCGAGA 701
Db 3269 AGGGATCTTGGCAGCGCGTGGCGGATCTCTCATTTGACTCTCTCCAAAGATGACCTCGAGA 3328
Qy 702 AAGTCAAGGCACCTTCAGAGAGAAATTGCGCCCTCTGAAAGAGAACGTGAGCCACGCTCAATG 761
Db 3329 AAGTCAAGGCACCTTCAGAGAGAAATTGCGCCCTCTGAAAGAGAACGTGAGCCACGCTCAATG 3388
Qy 762 ACCTTGTCTGCCAGCTTACCCTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCACTC 821
Db 3389 ACCTTGTCTGCCAGCTTACCCTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCACTC 3448
Qy 822 TGGAGAGACCTGAAACACAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCAAGC 881
Db 3449 TGGAGAGACCTGAAACACAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCAAGC 3508
Qy 882 AGCTGCATGAAGCCCAAGGAGCTTTGTCAGCATCTCAGCACTTTCTTTCCAGCTCTG 941
Db 3509 AGCTGCATGAAGCCCAAGGAGCTTTGTCAGCATCTCAGCACTTTCTTTCCAGCTCTG 3568
Qy 942 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTCTATATCAACACG 1001
Db 3569 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTCTATATCAACACG 3628
Qy 1002 AGACTCAAAACAACTTGTGGGACCATCCAAATATGACAGAGCTCTACAGTCTTTAGCTG 1061
Db 3629 AGACTCAAAACAACTTGTGGGACCATCCAAATATGACAGAGCTCTACAGTCTTTAGCTG 3688
Qy 1062 ACCTGAATAATGTGAGTCTCAGCTTATAGGAGTCCATGAAGTCCGAAAGACTGAGAG 1121
Db 3689 ACCTGAATAATGTGAGTCTCAGCTTATAGGAGTCCATGAAGTCCGAAAGACTGAGAG 3748
Qy 1122 AGGCCCTTTGTGGATCTCTTGAAGCTGTGAGCTGTGATGCTTGGACGAGCACA 1181
Db 3749 AGGCCCTTTGTGGATCTCTTGAAGCTGTGAGCTGTGATGCTTGGACGAGCACA 3808
Qy 1182 ACCTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTAATTAATTTGTTGACCACTA 1241
Db 3809 ACCTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTAATTAATTTGTTGACCACTA 3868
```

QY	1242	TTTATGACCCCTGGAGCAGACGACAA	CAATTGGTCAACGTCCCTCTCTCGTGGATA	1301
DB	3869	TTTATGACCCCTGGAGCAGACGACAA	CAAAATTTGGTCAACGTCCCTCTCTCGTGGATA	3928
QY	1302	TGTCGTGAAGTGGCTGCTGAATGTT	TATGATACGGGACGAAACAGGAGGATCCCGTGTCC	1361
DB	3929	TGTCGTGAAGTGGCTGCTGAATGTT	TATGATACGGGACGAAACAGGAGGATCCCGTGTCC	3988
QY	1362	TGTCCTTTTAAACTGGCATCATTTCC	TCTGTGTAAAGCACATTTGGGAAGCAAGTACAGAT	1421
DB	3989	TGTCCTTTTAAACTGGCATCATTTCC	TCTGTGTAAAGCACATTTGGGAAGCAAGTACAGAT	4048
QY	1422	ACCTTTTCAAGCAAGTGGGCAAGTTC	ACACAGGATTTGTGACCAAGCGAGCTGGGCTCC	1481
DB	4049	ACCTTTTCAAGCAAGTGGGCAAGTTC	ACACAGGATTTGTGACCAAGCGAGCTGGGCTCC	4108
QY	1482	TTCTGCATGATTCATCCAA	1501	
DB	4109	TTCTGCATGATTCATCCAA	4128	
RESULT 11				
ID	AAD37256	standard; DNA; 4966 BP.		
XX	AC	AAD37256;		
XX	DT	21-AUG-2002 (first entry)		
XX	DE	Adeno-associated virus vector plasmid, AAV-MCK-delta3990.		
XX	KW	Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;		
KW	KW	adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;		
KW	KW	Becker muscular dystrophy; ds.		
XX	OS	HomO sapiens.		
OS	OS	Unidentified.		
OS	OS	Chimeric.		
XX	PN	WO200183695-A2.		
XX	PD	08-NOV-2001.		
XX	PF	27-APR-2001; 2001WO-US013677.		
XX	PR	28-APR-2000; 2000US-0200777P.		
XX	PA	(XIAO/) XIAO X.		
XX	PI	Xiao X;		
XX	WPI	2002-049342/06.		
XX	PT	New dystrophin minigene for treating Duchenne or Becker muscular		
PT	PT	dystrophy comprises an N-terminal domain or modified N-terminal domain,		
PT	PT	rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin		
PT	PT	gene.		
XX	PS	Example 1; Page 59-60; 71pp; English.		
XX	CC	The present invention relates to an isolated nucleotide sequence encoding		
CC	CC	a dystrophin minigene. The minigene comprises N-terminal or modified N-		
CC	CC	terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4		
CC	CC	domains and cysteine-rich domains of dystrophin or utrophin genes. The		
CC	CC	invention also relates to a recombinant adeno-associated virus (AAV)		
CC	CC	comprising dystrophin minigene operably linked to an expression control		
CC	CC	element. The dystrophin minigene in operable linkage with an expression		
CC	CC	control element, in a recombinant adeno-associated virus or retrovirus is		
CC	CC	useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular		
CC	CC	dystrophy (BMD) in a mammalian subject. The present sequence is AAV		
CC	CC	vector plasmid construct containing human dystrophin minigenes, a muscle		
CC	CC	creatine kinase (MCK) promoter and a small polyA signal sequence		

XX	Sequence	4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 U; 0 Other;
SQ	Query Match	88.58; Score 1328.8; DB 6; Length 4966;
	Best Local Similarity	99.54; Pred. No. 0;
	Matches 1333; Conservative	0; Mismatches 7; Indels 0; Gaps 0;
QY	162	GGGTCTCTTTTACAAGACAGTCTTCAC
DB	2907	GGTCCCATTTTGAAGCCAGTTCGACCA
QY	222	TTCTGGTGTGGCTACAGCTGAAGATGAT
DB	2967	TTCTGGTGTGGCTACAGCTGAAGATGAT
QY	282	ACTTTCAGCAGTTCAGAAGCAGAACGAT
DB	3027	ACTTTCAGCAGTTCAGAAGCAGAACGAT
QY	342	CTAAGAAACCTGTAAATCATATGAGTA
DB	3087	CTAAGAAACCTGTAAATCATATGAGTA
QY	402	CTTTTGAAGAGTACAGAAACTCTACCA
DB	3147	CTTTTGAAGAGTACAGAAACTCTACCA
QY	462	CCCAGATGTCACTCGGCTTCTACAAAG
DB	3207	CCCAGATGTCACTCGGCTTCTACAAAG
QY	522	AATTGAACCTGCATCCCGCTGACTGGC
DB	3267	AATTGAACCTGCATCCCGCTGACTGGC
QY	582	AGGAACTTCAAGAGGCCACGGATGAGT
DB	3327	AGGAACTTCAAGAGGCCACGGATGAGT
QY	642	AGGGATCTCTGGCAGCCCGTGGCGAT
DB	3387	AGGGATCTCTGGCAGCCCGTGGCGAT
QY	702	AACTCAAGGCATTCGAGGAGAAATTT
DB	3447	AACTCAAGGCATTCGAGGAGAAATTT
QY	762	ACCTTGCTCGCCAGCTTACCACCTTT
DB	3507	ACCTTGCTCGCCAGCTTACCACCTTT
QY	822	TGGAGACCTGAACACACAGATGRRAG
DB	3567	TGGAGACCTGAACACACAGATGRRAG
QY	882	AGCTGATGAAGCCACACAGGACTTT
DB	3627	AGCTGATGAAGCCACACAGGACTTT
QY	942	TCAGGGTCCCTGGGAGAGAGCCATCT
DB	3687	TCAGGGTCCCTGGGAGAGAGCCATCT
QY	1002	AGACTCAAAACAACTTCTCTGGGACC
DB	3747	AGACTCAAAACAACTTCTCTGGGACC
QY	1062	ACCTGAATTAATGTACAGATTCAGCT
DB	3807	ACCTGAATTAATGTACAGATTCAGCT
QY	1122	AGGCCCTTTGCTGGATCTCTTGAGCT

QY 1062 ACCTGATATGTCAGATTCTCAGCTTATAGGACTGCCATGAACCTCCGAAGACTGCAGA 1121
 Db |||||
 QY 3931 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACCTCCGAAGACTGCAGA 3890
 Db |||||
 QY 1122 AGGCCCTTTGCTGGATCTCTGAGCTGTGTCAGCTGTCATGTGATGCTTGGACACACACA 1181
 Db |||||
 QY 3931 AGGCCCTTTGCTGGATCTCTGAGCTGTGTCAGCTGTCATGTGATGCTTGGACACACACA 3950
 Db |||||
 QY 1182 ACCTCAAGCAAAATGACACGCCATCGATATCTCGAGATTATTAATTTGTTGACCACTA 1241
 Db |||||
 QY 3951 ACCTCAAGCAAAATGACACGCCATCGATATCTCGAGATTATTAATTTGTTGACCACTA 4010
 Db |||||
 QY 1242 TTTATGACCGCTGGAGCAAGACACAAATTTGGTCAACGTCCTCTCTGCGTGATA 1301
 Db |||||
 QY 4011 TTTATGACCGCTGGAGCAAGACACAAATTTGGTCAACGTCCTCTCTGCGTGATA 4070
 Db |||||
 QY 1302 TGTGCTGAATGCTGCTGCTGAAATGTTTATGATACGGACGAAACAGGAGGATCCGTGTC 1361
 Db |||||
 QY 4071 TGTGCTGAATGCTGCTGCTGAAATGTTTATGATACGGACGAAACAGGAGGATCCGTGTC 4130
 Db |||||
 QY 1362 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAAGACAAAGTACAGAT 1421
 Db |||||
 QY 4131 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAAGACAAAGTACAGAT 4190
 Db |||||
 QY 1422 ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGATTTTGTGACAGCGGACGCTGGCTTC 1481
 Db |||||
 QY 4191 ACCTTTTCAGCAAGTGGCAAGTTTCAACAGATTTTGTGACAGCGGACGCTGGCTTC 4250
 Db |||||
 QY 1482 TTCTGTCATGATTTCTATCCAA 1501
 Db |||||
 QY 4251 TTCTGTCATGATTTCTATCCAA 4270
 Db |||||

RESULT 13
 RAD37264
 ID AD37264 standard; DNA; 5060 BP.
 AC AD37264;
 XX
 DT 21-AUG-2002 (first entry)
 DE Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.
 XX
 KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
 KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
 KW Becker muscular dystrophy; ds.
 XX
 OS Homo sapiens.
 OS Cytomegalovirus.
 OS Unidentified.
 OS Chimeric.
 XX
 FN WO200183695-A2.
 XX
 PD 08-NOV-2001.
 XX
 PF 27-APR-2001; 2001WO-US013677.
 XX
 PR 28-APR-2000; 2000US-0200777P.
 XX
 XX (XIAO/) XIAO X.
 XX
 XX Xiao X;
 XX
 XX WPI; 2002-049342/06.
 XX
 XX New dystrophin minigene for treating Duchenne or Becker muscular
 PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
 PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
 PT gene.
 XX
 XX Example 1; Page 70-71; 71pp; English.

CC The present invention relates to an isolated nucleotide sequence encoding
 CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
 CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
 CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
 CC invention also relates to a recombinant adeno-associated virus (AAV)
 CC comprising dystrophin minigene operably linked to an expression control
 CC element. The dystrophin minigene in operable linkage with an expression
 CC control element, in a recombinant adeno-associated virus or retrovirus is
 CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
 CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
 CC vector plasmid construct containing human dystrophin minigenes, a muscle
 CC creatine kinase (CK) enhancer, a cytomegalovirus (CMV) promoter and a
 CC small polyA signal sequence
 XX
 SQ Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 U; 0 Other;
 Query Match 88.5%; Score 1328.8; DB 6; Length 5060;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 162 GGGTCTTTTACAGACAGTCTTGACAGTGGAGCGTCTGCACCTTCTCTGAGGAAAC 221
 Db |||||
 QY 3001 GGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAGCGTCTGCACCTTCTCTGAGGAAAC 3060
 Db |||||
 QY 222 TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGAGCACCTATTGAGGCG 281
 Db |||||
 QY 3061 TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGAGCACCTATTGAGGCG 3120
 Db |||||
 QY 282 ACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAA 341
 Db |||||
 QY 3121 ACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAA 3180
 Db |||||
 QY 342 CTAAAGAACCTGTATCATGTAGTACTCTTGAGACTGTACGAATATTTCTGCACAGCAGC 401
 Db |||||
 QY 3181 CTAAAGAACCTGTATCATGTAGTACTCTTGAGACTGTACGAATATTTCTGCACAGCAGC 3240
 Db |||||
 QY 402 CTTTGGAAAGGACTAGAGAAACTCTTACCAGAGAGCCAGAGAGCTCCCTCTGAGAGAGAG 461
 Db |||||
 QY 3241 CTTTGGAAAGGACTAGAGAAACTCTTACCAGAGAGCCAGAGAGCTCCCTCTGAGAGAGAG 3300
 Db |||||
 QY 462 CCCAGATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTCAGTGGGAAA 521
 Db |||||
 QY 3301 CCCAGATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTCAGTGGGAAA 3360
 Db |||||
 QY 522 AATTGAACTGTCACCTCCCTGACTGGCAGAGAAAAAATAGATGAGACCCCTTGAAGACTCC 581
 Db |||||
 QY 3361 AATTGAACTGTCACCTCCCTGACTGGCAGAGAAAAAATAGATGAGACCCCTTGAAGACTCC 3420
 Db |||||
 QY 582 AGGAACTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTCGGCCGATCTCTCAAGATCACTCCGAGA 641
 Db |||||
 QY 3421 AGGAACTTCAAGAGGCCACGGATGAGTGGACCTCAAGCTCGGCCGATCTCTCAAGATCACTCC 3480
 Db |||||
 QY 642 AGGATCTCTGGCAGCCCGTGGCGGATCTCTCAATGACTCTCTCAAGATCACTCCGAGA 701
 Db |||||
 QY 3481 AGGATCTCTGGCAGCCCGTGGCGGATCTCTCAATGACTCTCTCAAGATCACTCCGAGA 3540
 Db |||||
 QY 702 AAGTCAAGGCACTTCGAGGAGAAATTTGCGCTCTTGAAGAGAACCGTGAGCCACGTCATG 761
 Db |||||
 QY 3541 AAGTCAAGGCACTTCGAGGAGAAATTTGCGCTCTTGAAGAGAACCGTGAGCCACGTCATG 3600
 Db |||||
 QY 762 ACCTTGCTCGCCAGCTTACCCTTTGGCATTTCAGCTCTCACCGTATACCTTCAGCACTC 821
 Db |||||
 QY 3601 ACCTTGCTCGCCAGCTTACCCTTTGGCATTTCAGCTCTCACCGTATACCTTCAGCACTC 3660
 Db |||||
 QY 822 TGAAGACCTGAAACACACAGATGGAAGCTTTCTGAGGTGGCCGTGAGGACCGAGTCAGGC 881
 Db |||||
 QY 3661 TGAAGACCTGAAACACACAGATGGAAGCTTTCTGAGGTGGCCGTGAGGACCGAGTCAGGC 3720
 Db |||||
 QY 882 AGCTGATGAAGCCACACAGGACTTTGGTCCAGCATCTCAGACCTTCTTTCCAGCTCTG 941
 Db |||||
 QY 3721 AGCTGATGAAGCCACACAGGACTTTGGTCCAGCATCTCAGACCTTCTTTCCAGCTCTG 3780
 Db |||||
 QY 942 TCCAGGGTCCCTGGGAGAGAGCCATCTGCCAAAACAAAGTGCCTTACTATATCAACACAG 1001
 Db |||||

QY 882 AGCTGCATGAGCCACAGGACCTTTGGTCCAGCATCTCAGACCTTTCTTTCCAGCTCTG 941
Db |||||
QY 3810 AGCTGCATGAGCCACAGGACCTTTGGTCCAGCATCTCAGACCTTTCTTTCCAGCTCTG 3869
Db |||||
QY 942 TCCAGGGTCCCTGGGAGAGCCATCTCGCCAAACAAAGTCCCTACTATATCAACACG 1001
Db |||||
QY 3870 TCCAGGGTCCCTGGGAGAGCCATCTCGCCAAACAAAGTCCCTACTATATCAACACG 3929
Db |||||
QY 1002 AGACTCAACAACTTGTGGGACCATCCAAAATGACAGAGCTCTACCACTTTAGCTG 1061
Db |||||
QY 3930 AGACTCAACAACTTGTGGGACCATCCAAAATGACAGAGCTCTACCACTTTAGCTG 3989
Db |||||
QY 1062 ACCTGAATAAGTCAGATCTTCAGCTATAGGACTGCCATGAAGTCCGAGAGCTGAGA 1121
Db |||||
QY 3990 ACCTGAATAAGTCAGATCTTCAGCTATAGGACTGCCATGAAGTCCGAGAGCTGAGA 4049
Db |||||
QY 1122 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCAAC 1181
Db |||||
QY 4050 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCAAC 4109
Db |||||
QY 1182 ACCTCAACCAAAATGACAGCCCATGGATATCTTCGAGATTAATAATGTTGACCACTA 1241
Db |||||
QY 4110 ACCTCAACCAAAATGACAGCCCATGGATATCTTCGAGATTAATAATGTTGACCACTA 4169
Db |||||
QY 1242 TTTATGACCGCTGGAGCAAGACACAAATTTGGTCAACGTCCTCTCTGCGTGATA 1301
Db |||||
QY 4170 TTTATGACCGCTGGAGCAAGACACAAATTTGGTCAACGTCCTCTCTGCGTGATA 4229
Db |||||
QY 1302 TGTGCTGAATGCTGCTGTAATTTATGATACGGGACGAAACAGGAGGATCCGTGTCC 1361
Db |||||
QY 4230 TGTGCTGAATGCTGCTGTAATTTATGATACGGGACGAAACAGGAGGATCCGTGTCC 4289
Db |||||
QY 1362 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAGTACAGAT 1421
Db |||||
QY 4290 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAGTACAGAT 4349
Db |||||
QY 1422 ACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCAAGCGAGGCTGGGCTCC 1481
Db |||||
QY 4350 ACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCAAGCGAGGCTGGGCTCC 4409
Db |||||
QY 1482 TTTGTCATGATCTCATCCAA 1501
Db |||||
QY 4410 TTTGTCATGATCTCATCCAA 4429
Db |||||

RESULT 15

ID ABK81998
XX ABK81998 standard; DNA; 5339 BP.
AC ABK81998;
XX
XX
XX 13-AUG-2002 (first entry)
XX
XX DNA encoding mini-dystrophin protein deltaR2-R21.
XX
XX Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
XX Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
XX Homo sapiens.
OS
XX Synthetic.
XX
XX W0200229056-A2.
PN
XX
XX 11-APR-2002.
XX
XX 04-OCT-2001; 2001WO-US031126.
XX
XX 06-OCT-2000; 2000US-0238848P.
XX
XX (UNMI) UNIV MICHIGAN.
PA
XX Chamberlain JS, Harper SQ;
PI
XX

WPI; 2002-435334/46.

XX A composition for preparing therapeutic drugs, has a mini-dystrophin peptide comprising a specific number of spectrin-like repeat domains, or a nucleic acid sequence encoding the mini-dystrophin peptide.

XX Example 6; Fig 13; 145pp; English.

XX The invention describes a composition comprising a mini-dystrophin peptide comprising a spectrin-like repeat domain, where the domain comprises n spectrin-like repeats, and contains no more than n spectrin-like repeats, where n is an even number between 4-24, or a nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the polynucleotide encoding it is useful as a medicament, for preparing a drug for therapeutic application and in the preparation of a composition for treatment of muscle disease, e.g. Duchenne's muscular dystrophy (DMD). This sequence represents a mini-dystrophin sequence of the invention

SQ Sequence 5339 BP; 1638 A; 1191 C; 1187 G; 1323 T; 0 U; 0 Other;

Query Match 88.5%; Score 1328.8; DB 6; Length 5339;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 162 GGGTCTTTTACAGACAGTTCTGACCAAGTGGAGCGTCTGCACCTTTCTTCGAGGAAC 221

Db |||||

QY 1875 GGTCCATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTTCGAGGAAC 1934

Db |||||

QY 222 TTTGCTGTGCTCAGCTGAAAGATGATGAATTAAGCCGACGACCTATTTCGAGGCG 281

Db |||||

QY 1935 TTTGCTGTGCTCAGCTGAAAGATGATGAATTAAGCCGACGACCTATTTCGAGGCG 1994

Db |||||

QY 282 ACTTTCAGAGTTCAGAGCAGAACGATGTACATAGGCTTCAAGAGGGAATTTGAAA 341

Db |||||

QY 1995 ACTTTCAGAGTTCAGAGCAGAACGATGTACATAGGCTTCAAGAGGGAATTTGAAA 2054

Db |||||

QY 342 CTAAGACCTGTAACTCATGACTACTCTTGAGACTGTACGAATATTTCTCAGAGAGCAGC 401

Db |||||

QY 2055 CTAAGACCTGTAACTCATGACTACTCTTGAGACTGTACGAATATTTCTCAGAGAGCAGC 2114

Db |||||

QY 402 CTTTGGAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCCTCTCAGAGAGAGAG 461

Db |||||

QY 2115 CTTTGGAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCCTCTCAGAGAGAGAG 2174

Db |||||

QY 462 CCCAGATGTCACTCGGCTTCTAGAAAAGCAGCGCTGAGGAGTCAATACTGAGTGGGAAA 521

Db |||||

QY 2175 CCCAGATGTCACTCGGCTTCTAGAAAAGCAGCGCTGAGGAGTCAATACTGAGTGGGAAA 2234

Db |||||

QY 522 AATTGAACCTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCC 581

Db |||||

QY 2235 AATTGAACCTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCC 2294

Db |||||

QY 582 AGGAATCTTAAGAGGCCACGATGAGCTGAACTCAAGCTGCGCCAAAGCTGAGGTGATCA 641

Db |||||

QY 2295 AGGAATCTTAAGAGGCCACGATGAGCTGAACTCAAGCTGCGCCAAAGCTGAGGTGATCA 2354

Db |||||

QY 642 AGGATCTTGGCAGCCGCTGGGCGATCTCTCTTGAAGTCTCTCCAGATCAGCTCGAGA 701

Db |||||

QY 2355 AGGATCTTGGCAGCCGCTGGGCGATCTCTCTTGAAGTCTCTCCAGATCAGCTCGAGA 2414

Db |||||

QY 702 AAGTCAAGGCACTTCGAGGAGAAATTTGCGCTCTGAAAAGAGAACGTCGAGCCAGTCAATG 761

Db |||||

QY 2415 AAGTCAAGGCACTTCGAGGAGAAATTTGCGCTCTGAAAAGAGAACGTCGAGCCAGTCAATG 2474

Db |||||

QY 762 ACCTTGCTCCAGCTTACCACTTTGGGCACTCAGCTCTCAGCTATACCTTCAGCACTC 821

Db |||||

QY 2475 ACCTTGCTCCAGCTTACCACTTTGGGCACTCAGCTCTCAGCTATACCTTCAGCACTC 2534

Db |||||

QY 822 TGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCCCTCGAGGAGCCAGTCAGGC 881

Db |||||

QY 2535 TGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCCCTCGAGGAGCCAGTCAGGC 2594

Db |||||

QY 882 AGTGCATGAAGCCGACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTCTG 941
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2595 AGTGCATGAAGCCGACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTCTG 2654
QY 942 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCAG 1001
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2655 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCAG 2714
QY 1002 AGACTCAAAACAATTGCTGGGAGAGCCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTG 1061
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2715 AGACTCAAAACAATTGCTGGGAGAGCCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTG 2774
QY 1062 ACCTGAATATGTCAGATCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAGA 1121
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2775 ACCTGAATATGTCAGATCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAGA 2834
QY 1122 AGGCCCTTTGCTTGGATCTCTTGAAGCTGTGAGCTGTGATGCTGTGATGCTTGGACCAACA 1181
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2835 AGGCCCTTTGCTTGGATCTCTTGAAGCTGTGAGCTGTGATGCTGTGATGCTTGGACCAACA 2894
QY 1182 ACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCACTA 1241
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2895 ACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCACTA 2954
QY 1242 TTTATGACCCCTGGAGCAAGAGACAAACAATTGGTCAACGTCCTCTCTGCGTGGATA 1301
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2955 TTTATGACCCCTGGAGCAAGAGACAAACAATTGGTCAACGTCCTCTCTGCGTGGATA 3014
QY 1302 TGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTC 1361
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3015 TGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTC 3074
QY 1362 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAAGCAAGTACAGAT 1421
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3075 TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAAGCAAGTACAGAT 3134
QY 1422 ACCTTTTCAAGCAAGTGGCAAGTTCAAACAGATTTTGTGACCAAGGAGGCTGGGCTCC 1481
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3135 ACCTTTTCAAGCAAGTGGCAAGTTCAAACAGATTTTGTGACCAAGGAGGCTGGGCTCC 3194
QY 1482 TTCGATGATTTCTATCCAA 1501
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3195 TTCGATGATTTCTATCCAA 3214

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 19, 2004, 03:01:40 ; Search time 77,4052 Seconds
(without alignments)
10761.315 Million cell updates/sec

Title: US-09-845-416-12_COPY_1500_3000

Perfect score: 1501

Sequence: 1 agaagatctagaacaagaac.....ttctgcatgattctatccaa 1501

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 27747546 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1328.8	88.5	5952	4	US-09-687-875A-1
2	1317.8	87.8	13977	4	US-09-484-970B-60
3	1149.6	76.6	13207	3	US-08-836-022A-10
4	1149.6	76.6	13207	3	US-08-427-048A-10
5	547.4	36.5	6045	4	US-09-091-501B-7
6	547.4	36.5	10320	4	US-09-091-501B-9
7	517.8	34.5	3915	4	US-09-976-594-93
8	69.8	4.7	200	4	US-09-091-501B-5
9	69	4.6	200	4	US-09-091-501B-4
10	65	4.3	200	4	US-09-091-501B-6
11	53.4	3.6	7218	1	US-08-233-463-44
12	44.2	2.9	2574	4	US-09-668-313A-10
13	42.8	2.9	1690	4	US-09-620-312D-69
14	42.8	2.9	7812	3	US-09-368-590-1
15	38.6	2.6	1995	1	US-08-425-069-3
16	38.6	2.6	1995	2	US-08-317-844B-3
17	38.4	2.6	7672	4	US-09-220-132-24
18	38.2	2.5	428	4	US-09-668-313A-3
19	38.2	2.5	4439	4	US-09-668-313A-17
20	34.8	2.3	1830121	4	US-09-557-884-1
21	34.8	2.3	1830121	4	US-09-643-990A-1
22	34.6	2.3	474	4	US-09-621-976-18033
23	34.6	2.3	2277	1	US-08-676-947-5
24	34.6	2.3	2277	1	US-08-676-947-5
25	34.6	2.3	2277	2	US-09-098-487-5
26	34.4	2.3	1047	4	US-09-671-950-1
27	34.4	2.3	1047	4	US-09-671-950-3

28 34.4 2.3 1047 4 US-09-671-950-5 Sequence 5, Appli
29 34.4 2.3 1047 4 US-09-671-950-7 Sequence 7, Appli
30 34.4 2.3 1047 4 US-09-671-950-9 Sequence 9, Appli
31 34.4 2.3 1047 4 US-09-671-950-11 Sequence 11, Appli
32 34.4 2.3 1047 4 US-09-671-950-13 Sequence 13, Appli
33 34.2 2.3 750 4 US-08-961-527-370 Sequence 370, Appl
34 34.2 2.3 1620 4 US-08-858-207A-56 Sequence 56, Appl
35 34.2 2.3 2800 2 US-08-874-138-1 Sequence 1, Appli
36 34.2 2.3 2800 3 US-08-874-138-5 Sequence 5, Appli
37 34.2 2.3 2800 3 US-08-879-941-1 Sequence 1, Appli
38 34.2 2.3 2800 3 US-08-879-941-3 Sequence 3, Appli
39 34.2 2.3 2800 4 US-09-747-116-1 Sequence 1, Appli
40 34.2 2.3 2800 4 US-09-747-116-3 Sequence 3, Appli
41 34.2 2.3 32768 4 US-08-961-527-71 Sequence 71, Appl
42 34 2.3 648 4 US-09-252-991A-10033 Sequence 10033, A
43 34 2.3 762 4 US-09-252-991A-9821 Sequence 9821, Ap
44 34 2.3 1844 4 US-09-252-991A-10161 Sequence 10161, A
45 34 2.3 2235 3 US-09-153-804-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-687-875A-1

; Sequence 1, Application US/09687875A

; Patent No. 6544786

; GENERAL INFORMATION:

; APPLICANT: Xiao, Xiao

; APPLICANT: Liu, Paul

; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED

; FILE REFERENCE: 00792

; CURRENT APPLICATION NUMBER: US/09/687,875A

; CURRENT FILING DATE: 2000-10-13

; PRIOR APPLICATION NUMBER: 60/158,868

; PRIOR FILING DATE: 1999-10-15

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 1

; LENGTH: 5952

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (2897)..(2898)

; OTHER INFORMATION: S4 junction site

; NAME/KEY: misc feature

; LOCATION: (3198)..(3199)

; OTHER INFORMATION: S2 junction site

US-09-687-875A-1

Query Match 88.5%; Score 1328.8; DB 4; Length 5952;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 162 GGGTCTTTTACAGACAGTCTTGACAGTGAAGGCTGACACCTTTCTTGACAGAAC 221

Db 3284 GGTCCATTGGAGCCAGTTCTGACAGTGAAGGCTGACACCTTTCTTGACAGAAC 3343

QY 222 TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGAGGACCTATTGGAGCG 281

Db 3344 TTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGAGGACCTATTGGAGCG 3403

QY 282 ACTTCCAGAGTTCAGAGCAGAGATGATACATAGGCGCTTCAAGAGGATTCAGAAA 341

Db 3404 ACTTCCAGAGTTCAGAGCAGAGATGATACATAGGCGCTTCAAGAGGATTCAGAAA 3463

QY 342 CTAAGAACCTGTAATCATGAGTACCTTTGAGACTGTAGAAATATTTCTGACAGACAGC 401

Db 3464 CTAAGAACCTGTAATCATGAGTACCTTTGAGACTGTAGAAATATTTCTGACAGACAGC 3523

QY 402 CTTTGAAGACTGTAGAGAACTCTACAGAGCCCGAGAGCTGCTCTCTGAGGAGAG 461

Db 402 CTTTGAAGACTGTAGAGAACTCTACAGAGCCCGAGAGCTGCTCTCTGAGGAGAG 461

Db 9198 ACCTTGCTGGCAGCTTACCACTTTGGCATTGAGCTTACCGTATAAAGCTCAGCACTC 9257
Qy 822 TGGAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCGTGCAGAGCGAGTCAAGC 881
Db 9258 TGGAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCGTGCAGAGCGAGTCAAGC 9317
Qy 882 AGCTGCATGAAGCCACACAGGACCTTTGGTCCAGATCTCAGCACTTTCTTTCCAGCTCTG 941
Db 9318 AGCTGCATGAAGCCACACAGGACCTTTGGTCCAGATCTCAGCACTTTCTTTCCAGCTCTG 9377
Qy 942 TCCAGGCTCCCTGGAGAGACCACTCTGCCAACAAGTGCCTTACTATATCAACACG 1001
Db 9378 TCCAGGCTCCCTGGAGAGACCACTCTGCCAACAAGTGCCTTACTATATCAACACG 9437
Qy 1002 AGACTCAAAACAACTTGCTGGACCACTCCAAATGACAGACTCTACAGCTCTTTAGCTG 1061
Db 9438 AGACTCAAAACAACTTGCTGGACCACTCCAAATGACAGACTCTACAGCTCTTTAGCTG 9497
Qy 1062 ACCTGAATAATGTCAGATTTCTAGCTTATAGACTGTCATGAAGTCCGAAAGTCAAGTCAAG 1121
Db 9498 ACCTGAATAATGTCAGATTTCTAGCTTATAGACTGTCATGAAGTCCGAAAGTCAAGTCAAG 9557
Qy 1122 AGGCCCTTGCTGGATCTTGAGCTCTGAGCTGTCAGCTGTCATGTCCTTGAGCCAGCACA 1181
Db 9558 AGGCCCTTGCTGGATCTTGAGCTCTGAGCTGTCAGCTGTCATGTCCTTGAGCCAGCACA 9617
Qy 1182 ACCTCAAGCAAAATGACAGCCCATGATATCTGAGATTAATTAATTTGTTGACCACTA 1241
Db 9618 ACCTCAAGCAAAATGACAGCCCATGATATCTGAGATTAATTAATTTGTTGACCACTA 9677
Qy 1242 TTTATGACCGCTGGAGAGACGACACAACTTTGTCAGCTCCCTCTCTGGTGGATA 1301
Db 9678 TTTATGACCGCTGGAGAGACGACACAACTTTGTCAGCTCCCTCTCTGGTGGATA 9737
Qy 1302 TGTGCTGAATGCTGCTGAATGTTATGATAGCGGACGACAAAGGAGGATCGTCTCC 1361
Db 9738 TGTGCTGAATGCTGCTGAATGTTATGATAGCGGACGACAAAGGAGGATCGTCTCC 9797
Qy 1362 TGTCTTTAAACTGGGATATTTCCCTGTTAAAGCAGATTTGGAAGCAAGTACAGAT 1421
Db 9798 TGTCTTTAAACTGGGATATTTCCCTGTTAAAGCAGATTTGGAAGCAAGTACAGAT 9857
Qy 1422 ACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGGAGCTTGGGCTC 1480
Db 9858 ACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGGAGCTTGGGCTC 9917
Qy 1481 CTTCTGCATGATTTATCCAA 1501
Db 9918 CTTCTGCATGATTTATCCAA 9938

RESULT 3
US-08-836-022A-10/c
; Sequence 10, Application US/08836022A
; Patent No. 6001557
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Fisher, Krishna J.
; APPLICANT: Chen, Shu-Jen
; APPLICANT: Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,022A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/331,381
; FILING DATE: 28-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNPVN.008PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19307 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; US-08-836-022A-10

Query Match 76.6%; Score 1149.6; DB 3; Length 19307;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 1221; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

Qy 162 GGTTCTTTTACAAGACAGTCTGACCACTGGAAGCGTCTGCACCTTTCTTCTGAGGAAC 221
Db 6096 GGTTCCATTTGGAAGCAAGTTCTGACCACTGGAAGCGTCTGATCTTTCTTCTGAGGAAC 6037
Qy 222 TTTCTGTGTGGCTACAGCTGGAAGATGATGAATTAAGCCGGCAGGACCACTTATGGAGCG 281
Db 6036 TTTCTGTGTGGCTACAGCTGGAAGATGATGAATTAAGCCGGCAGGACCACTTATGGAGCG 5977
Qy 282 ACTTCCACAGTTTCAAGACGACAGTGTACATAGGGCTTCAAGAGGGAATTGAAA 341
Db 5976 ATTTCACAGCTTCAAGACGACAGTGTACATAGGGCTTCAAGAGGGAATTGAAA 5917
Qy 342 CTAAAGAACTGTATCATGAGTACTCTTGAAGCTGTACGAATATTTCTGACAGAGCAGC 401
Db 5916 CTAAAGAACTGTATCATGAGTACTCTGAGACTGTGAGATATTTCTGACAGAGCAGC 5857
Qy 402 CTTTGAAGAGCTAGAGAACTCTACAGAGCCGACAGAGCTGCTCTCTGAGGAGAGAG 461
Db 5856 CTTTGAAGAGCTAGAGAACTCTACAGAGCCGACAGAGCTGCTCTCTGAGGAGAGAG 5797
Qy 462 CCCAGAACTGCACTCGGCTTCTACGAAAGCAGGCTGAGGAGCTCAATCTGAGTGGAAA 521
Db 5796 CTGAGAACTGCACTCGGCTTCTACGAAAGCAGGCTGAGGAGCTCAACGCTGATGGACA 5737
Qy 522 AATTGAACCTGCACTCGGCTGAGTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC 581
Db 5736 AATTGAACCTGCGCTCAGCTGATTTGACAGAGAAAATAGATGAGACTCTTTGAAAGACTCC 5677
Qy 582 AGGAACCTCAAGAGCCAGGATGAGTGGAGCTCAAGCTGCGCAAGCTGAGGTGATCA 641
Db 5676 AGGAACCTCAAGAGCTGCGGATGAGTGGAGCTCAAGCTGCGCAAGCTGAGGTGATCA 5617
Qy 642 AGGGATCTCGGAGCCCGTGGGCGATCTCTCTCATTTGACTCTCTCCAAGATCACTCGAGA 701
Db 5616 AGGGATCTCGGAGCCAGTGGGCGATCTCTCTCATTTGACTCTCTGCAAGATCACTTGAAA 5557
Qy 702 AGTCAAGGACTTTCAGAGAGAAATGCGCTCTGAAGAGAACTGAGCAGCCAGCTCAATG 761
Db 5556 AAGTCAAGGACTTTCAGAGAGAAATGCGCTCTGAAGAGAAATGAGTCAATCTGTCATG 5497
Qy 762 ACCTTCTCGCCAGCTTACACTTTGGGCACTTCAAGCTCTCAGCTTATACCTCAGCACTC 821
Db 5496 ACCTTCTCGCACTCAGCTGACCACTGCGGCACTTCAAGCTCTCAGCTTATACCTCAGCACTT 5437

882 TGGAGAGCTGAACACCGAGTGAAGCTTCTGCGAGGTGGCGTGCAGAGCCGAGTCAAGC 881
5436 TGGAGAGCTGAACACCGAGTGAAGCTTCTGCGAGGTGGCGTGCAGAGCCGAGTCAAGC 5377
882 AGCTGCATGAGCCACAGGAGCTTGGTCCAGATCTCAGACATCTTCTTCCAGCTG 941
5376 AGCTGCATGAGCCACAGGAGCTTGGTCCAGATCTCAGACATCTTCTTCCAGCTG 5317
942 TCCAGGTCCTGGGAGAGCCATCTCGCCAAACAAAGTCCCTACTATATCAACACG 1001
5316 TCCAGGTCCTGGGAGAGCCATCTCACCACAAAGTCCCTACTATATCAACACG 5257
1002 AGACTCAACAACTTGGTGGAGACCTCCAAATGACAGAGCTTACAGTCTTAGCTG 1061
5256 AGACCAACAACTTGGTGGAGACCTCCAAATGACAGAGCTTACAGTCTTAGCTG 5197
1062 ACCTGAATAATGTCAGATCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAGA 1121
5196 ACCTGAATAATGTCAGATCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAGA 5137
1122 AGGCCCTTGGTGGATCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAG 1181
5136 AGGCCCTTGGTGGATCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAG 5077
1182 ACCTCAAGCAAAATGACAGAGCCATGATATCTGCGAGATTATTAATTTTGGACCACTA 1241
5076 ACCTCAAGCAAAATGACAGAGCCATGATATCTGCGAGATTATTAATTTTGGACCACTA 5017
1242 TTATGACCGCTGGAGCAAGAGACACAAATTTGGTCAAGCTCCCTCTCTGGGTGATA 1301
5016 TTATGACCGCTGGAGCAAGAGACACAAATTTGGTCAAGCTCCCTCTCTGGGTGATA 4957
1302 TGTGCTGACTGCTGCTGATGATTTTATGATGCGGAGCAAGAGGAGGATCGGTGTC 1361
4956 TGTGCTGACTGCTGCTGATGATTTTATGATGCGGAGCAAGAGGAGGATCGGTGTC 4897
1362 TGTGCTGACTGCTGCTGATGATTTTATGATGCGGAGCAAGAGGAGGATCGGTGTC 1421
4896 TGTGCTGACTGCTGCTGATGATTTTATGATGCGGAGCAAGAGGAGGATCGGTGTC 4837
1422 ACCCTTTCAAGCAAGTGCAGATTCACAGAGATTTTGGAGCCAGCGAGGCTGGGCTCC 1481
4836 ACCCTTTCAAGCAAGTGCAGATTCACAGAGATTTTGGAGCCAGCGAGGCTGGGCTCC 4777
1482 TTCTGCATGATTCATCAAA 1501
4776 TTCTGCATGATTCATCAAA 4757

RESULT 4

US-09-427-048A-10/c
; Sequence 10, Application US/09427048A
; Patent No. 6203975

GENERAL INFORMATION:

APPLICANT: Trustees of the University of Pennsylvania

Wilson, James M.
Fisher, Krishna J.
Chen, Shu-Jen
Weitzman, Matthew

TITLE OF INVENTION: Improved Adenovirus and
Methods of Use Thereof

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Cntr, P O Box 457

CITY: Spring House

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,048A
FILING DATE: 21-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,022
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNPN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRAINEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-427-048A-10

Query Match 76.6%; Score 1149.6; DB 3; Length 19307;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 1221; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
QY 162 GGGTCTTTTCAAGACAGTCTTGACCAAGTGAAGCGTCTGACACCTTTCTTGACGAGAAC 221
Db 6096 GGTCCCATTTGAAGCAAGTCTGACCAAGTGAAGCGTCTGACACCTTTCTTGACGAGAAC 6037
QY 222 TTCTGGTGGCTAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGCG 281
Db 6036 TTCTGGTGGCTAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGCG 5977
QY 282 ACTTCCAGCAGTCTCAGAAAGCAGAACTGTACATAGGCGCTTCAAGAGGAAATTTGAAA 341
Db 5976 ATTTCCAGCAGTCTCAGAAAGCAGAACTGTACATAGGCGCTTCAAGAGGAAATTTGAAA 5917
QY 342 CTAAAGAACCTGTAATCATAGTACTCTTGAGACTGTGAGAAATTTCTGACAGAGCAGC 401
Db 5916 CTAAAGAACCTGTAATCATAGTACTCTTGAGACTGTGAGAAATTTCTGACAGAGCAGC 5857
QY 402 CTTTGGAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCTCTCTGAGGAGAGAG 461
Db 5856 CTTTGGAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCTCTCTGAGGAGAGAG 5797
QY 462 CCCAGAACTGCTACTCGGCTTTCTAGAAAGAGGCTGAGAGGTCAATCTGAGTGGGAAA 521
Db 5796 CTCAGAACTGCTACTCGGCTTTCTAGAAAGAGGCTGAGAGGTCAATCTGAGTGGGAAA 5737
QY 522 AATTGAACCTGCACCTCCGCTGACTGCGAGAGAAATAGATGAGACCTTGAAGAGCTCC 581
Db 5736 AATTGAACCTGCACCTCCGCTGACTGCGAGAGAAATAGATGAGACCTTGAAGAGCTCC 5677
QY 582 AGGAACTTCAAGAGGCCAGGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCA 641
Db 5676 AGGAACTTCAAGAGGCCAGGATGAGCTGGACCTCAAGCTCGGCCAAGCTGAGGTGATCA 5617
QY 642 AGGATCTCTGCGAGCCGCTGGCGATCTCTCATTTGACTCTCTCCAGATCACCTCGAGA 701
Db 5616 AGGATCTCTGCGAGCCGCTGGCGATCTCTCATTTGACTCTCTCCAGATCACCTCGAGA 5557
QY 702 AAGTCAAGGCACTTCCAGAGGAGAAATTTGGCGCTCTGAAAAGAGAACTGAGCCACGTCATG 761
Db 5556 AAGTCAAGGCACTTCCAGAGGAGAAATTTGGCGCTCTGAAAAGAGAACTGAGCCACGTCATG 5497
QY 762 ACCTTGCTGCGCAGCTTACCACTTTGGGCACTTACGCTCTACCGTATTAACCTCAGCACT 821
Db 5496 ACCTTGCTGCGCAGCTTACCACTTTGGGCACTTACGCTCTACCGTATTAACCTCAGCACT 5437

Qy	876	TCAGGCAGCTGCATGAAGCCACACAGGACCTTTGGTCAGCATCTCAGGACCTTTCTTTCCCA	935
Db	4110	TTAAACAGCTTCAGGAAGCCACAGAGATTTTGACCATCCTCTCAGCATTTTCTCTCTA	4169
Qy	936	CGTCTGTCAGGGTCCCTGGGAGAGACCATCTCGCCAAACAAAGTGCCTTACTATATCA	995
Db	4170	CGTCAGTCCAGCTGCCGTGGCAAGATCCATTTTTCACATAATAAAGTGCCTATTTACATCA	4229
Qy	996	ACACAGAGACTCAACACCACTTGCTGGAGCCATCCCAAAATGACAGAGCTCTACAGTCTT	1055
Db	4230	ACCATCAACACAGACCACCTGTTGGGACCATCTTAAATGACCGAACTCTTTCAAATCCC	4288
Qy	1056	TAGCTGACCTGAATTAATGTCCAGATCTCAGCTTATAGGACTGCCATGAAACTCCGAAGAC	1115
Db	4290	TTGTGACCTGAATAATGTACGTTTTTCTGCCTACCGTACAGCAATCAAAATCCGAAGAC	4349
Qy	1116	TGCAGAAGGCCCTTTGCTTGATCTCTTGAGCCGTGTCAGCTGCGATGTATGCCTTGAGCC	1175
Db	4350	TACAAAGAAGCACTATGTTTGGATCTCTTAGAGTTAGTACAACAATGAATTTTCAAC	4409
Qy	1176	AGCACAACTCTCAAGCAAAATGACGAGCCCATGGATATCTGCGAGATTAATTAATGTTGA	1235
Db	4410	AGCACAGTTGAACCAAAATGACAGGCTCCTCAGTGTCCAGATGTGCATCAACTGTCTGA	4469
Qy	1236	CCACTATTATGACGGCTGGAGCAAGACACAAATTTGGTCAACGTCCTCTCTCGG	1295
Db	4470	CAACAACTTATGATGGACTTGAGCAATGCATTAAGGACCTGGTCAACGTTCCACTCTGTG	4529
Qy	1296	TGGATATGTGCTGAACTGCGTCTGTAATGTTTATGATAGGGACGCAACAGGAGNATCC	1355
Db	4530	TTGATATGTGCTCAATTTGGTTGCTCAATGCTATGACACGGGTGCAACTTGGAAAAATTA	4589
Qy	1356	GTGTCCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTTAAAGCACATTTTGGAAAGACAAAT	1415
Db	4590	GAGTGCAGAGTCTGAAGATTGGATTAATGCTCTCTCTCCAAAGGTCTCTTGGAGAAAAAT	4649
Qy	1416	ACGATACCTTTTCAACCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGCAGCGCTGG	1475
Db	4650	ACAGATATCTCTTTAAGGAAGTTTGGGGGGCGACAGAAATGTGTGACCGAGGCGAGCTGG	4709
Qy	1476	GCCTCCTCTGCATGATTTCTATCCA	1500
Db	4710	GCCTGTACTTTCATGATGCCATCA	4734

RESULT 6

```

US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 10320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(10312)

```



```
QY 936 CGTCTGTCAGGTCCTGGGAGAGACCCATCTGCCCCAAACAAAGTGCCCTACTATATCA 995
Db 8445 CGTCAGTCCAGCTCCCGTGGCAAAAGATCCATTTTACATAATAAAGTGCCCTATTACATCA 8504
QY 996 ACCACGAGACTCAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCACTTT 1055
Db 8505 ACCATCAACACAGACCACTTGTGGGACCATCTCAAAATGACCGAACTCTTCAATCCC 8564
QY 1056 TAGCTGACCTGAATTAATGTGAGATCTCAGCTTATAGAGCTGCCATGAAGACTCCGAAGAC 1115
Db 8565 TTGCTGACCTGAATTAATGTGAGATCTTCTGCTTACCGGTACAGCAATCAAAATCCGAAGAC 8624
QY 1116 TGCAGAAGGCCCTTGTGCTTGGATCTTGTAGCCTGTGAGCTGCTGATGATGCTTGGACC 1175
Db 8625 TACAAAAGCACTATGTTGGATCTCTAGATTGAGTACAACAAAATGAAATTTTCAAC 8684
QY 1176 AGCAAACTCAACAAAATGACAGGCCCATGGATATCCTGCAGATTATTAATGTTTGA 1235
Db 8685 AGCAAGTTGAACCAAAATGACAGCTCCTCAGTGTTCAGATGTCATCAACTGCTGA 8744
QY 1236 CCATATTATGACGGCTGGAGCAAGACAAACAATTTGGTCAAGCTCCCTCTCTGCG 1295
Db 8745 CAACAACCTTATGATGGACTTGGAGCAATGCAATAAGGACCTGGTCAACGTTCCACTCTGTG 8804
QY 1296 TGGATATGCTGCTGAATGCTGCTGATGTTTATGATAGGGACGAACAGGAGGATCC 1355
Db 8805 TTGATATGCTGCTCAATTTGGTGTCTCAATGCTATGACAGGGTCTGAACTGGAAAATTA 8864
QY 1356 GTGCTCTGCTTTTAAACTGGCAATCTTCCCTGTGTAAAGCACTTTGGAAGACAAGT 1415
Db 8865 GAGTGCAGAGTCTGAAGATTGGATTAATGTCTCTCTCCAAAGTCTCTTGGAGAAAAT 8924
QY 1416 ACAGATACCTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGGACCGCAGGCTGG 1475
Db 8925 ACAGATATCTCTTAAAGGAAGTTGGGGCGGACAGAAAATGTGTGACAGGCGAGCTGG 8984
QY 1476 GCCTCTCTCTCATGATTTCTATCCA 1500
Db 8985 GCCTGTACTTCAATGATGCCATCCA 9009

RESULT 7
US-09-976-594-93
; Sequence 93, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 93
; LENGTH: 3915
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. 6673549 290344.1
US-09-976-594-93

Query Match 34.5%; Score 517.8; DB 4; Length 3915;
Best Local Similarity 61.5%; Pred. No. 96-158;
Matches 846; Conservative 0; Mismatches 527; Indels 2; Gaps 1;

QY 126 GATGGCAACATCTGTAGATGGACAGAGACCGCTGGGTCTTTTACAAGACAGATTCTG 185
Db 555 GTTGAATGAATAAATAAATAAAGTCTCAACACCTCCGGCTCGCCTAGAGGCTTCTCAG 614
```

Db 1693 GAGGATCTCTGTCACCTGTCCTGTGGACATGAGCTCAATGGCTCCTCAATG 1752
QY 1326 TTTATGATACGGACGACAGGAGGATCCGCTGCTCTCTTTTAAACTGGCATATT 1385
Db 1753 TTTTGTATAGTGTGCGAGCGGAAGATGCGGCAATGTCTTTTAAAGCTGGCAT 1812
QY 1386 CCTGTGTAAGCAATTGGAAGCAAGTACAGTACATCTTTTCAAGCAAGTGCAGTT 1445
Db 1813 GCTTGTGTGGCAGGCAAGTGAAGGAAAAAATTCAGTACCTCTTCAGCAAGTGCCTCACT 1872
QY 1446 CAACAGGATTTTGTACACGAGCGAGCTGGCCCTCTTCTCATGATTCATCA 1500
Db 1873 CAGGAGCAGTGTGACCGAGCGCCACCTTGTGTCTCTCTTCATGAGCCATTC 1927

RESULT 8

US-09-091-501B-5

; Sequence 5, Application US/09091501B

; Patent No. 6518413

; GENERAL INFORMATION:

; APPLICANT: Tinsley, Jonathon M

; APPLICANT: Davies, Kay E

; TITLE OF INVENTION: Utrrophin gene expression

; FILE REFERENCE: 620-42

; CURRENT APPLICATION NUMBER: US/09/091,501B

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: PCT/GB96/03156

; PRIOR FILING DATE: 1996-12-19

; PRIOR APPLICATION NUMBER: GB 9525962.8

; PRIOR FILING DATE: 1995-12-19

; PRIOR APPLICATION NUMBER: GB 9615797.9

; PRIOR FILING DATE: 1996-07-26

; PRIOR APPLICATION NUMBER: GB 9622174.2

; PRIOR FILING DATE: 1996-10-24

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 5

; LENGTH: 200

; TYPE: DNA

; ORGANISM: Rattus sp.

; ORGANISM: Homo sapiens

US-09-091-501B-5

Query Match 4.7%; Score 69.8; DB 4; Length 200;
Best Local Similarity 67.6%; Pred. No. 1.1e-12;
Matches 98; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 5 GATCTAGAACAGCAAGTCAAGTCAAGTCAATCTCTCACTCAGTGGTGGTGGTGGTGGT 64
Db 56 GACCTCGAAGTGTGAGCGAGGTAAGTGAATTCCTTAATCTCACTCAGTGGTGGTGGTGGT 115
QY 65 GAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGCAAACTTAAGTATTGGGAGAT 124
Db 116 GAACACAGTGGGAGAGCGCCACAGCTGTTTGAAGATCAGTTACAGAACTGGGTGAG 175
QY 125 CGATGGGCAAACTCTGTAGATGGA 149
Db 176 CGTGGACAGCTGTATGCCGTGGA 200

RESULT 9

US-09-091-501B-4

; Sequence 4, Application US/09091501B

; Patent No. 6518413

; GENERAL INFORMATION:

; APPLICANT: Tinsley, Jonathon M

; APPLICANT: Davies, Kay E

; TITLE OF INVENTION: Utrrophin gene expression

; FILE REFERENCE: 620-42

; CURRENT APPLICATION NUMBER: US/09/091,501B

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: PCT/GB96/03156

; PRIOR FILING DATE: 1996-12-19

; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Mus sp.
; ORGANISM: Homo sapiens
US-09-091-501B-4

Query Match 4.6%; Score 69; DB 4; Length 200;
Best Local Similarity 66.4%; Pred. No. 2.1e-12;
Matches 99; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 AGAAGATCTAGAACAGCAAGTCAAGTCAAGTCAATCTCTCACTCAGTGGTGGTGGTGGT 60
Db 52 AAATGACCTTGAAGCTGAACAGGTGAAGTAAATTCCTTAATCTCACTCAGTGGTGGTGGT 111
QY 61 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGCAAACTTAAGTATTGGG 120
Db 112 GGATGAACACAGTGGGAGAGTGCACAGCTCTTCTGGAAGATCAGTTACAGAACTGGG 171
QY 121 AGATCGATGGGCAAACTCTGTAGATGGA 149
Db 172 TGAGCGCTGGACAGCTGTATGCCGTGGA 200

RESULT 10

US-09-091-501B-6

; Sequence 6, Application US/09091501B

; Patent No. 6518413

; GENERAL INFORMATION:

; APPLICANT: Tinsley, Jonathon M

; APPLICANT: Davies, Kay E

; TITLE OF INVENTION: Utrrophin gene expression

; FILE REFERENCE: 620-42

; CURRENT APPLICATION NUMBER: US/09/091,501B

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: PCT/GB96/03156

; PRIOR FILING DATE: 1996-12-19

; PRIOR APPLICATION NUMBER: GB 9525962.8

; PRIOR FILING DATE: 1996-12-19

; PRIOR APPLICATION NUMBER: GB 9615797.9

; PRIOR FILING DATE: 1996-07-26

; PRIOR APPLICATION NUMBER: GB 9622174.2

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 200

; TYPE: DNA

; ORGANISM: Homo sapiens

; ORGANISM: Homo sapiens

US-09-091-501B-6

Query Match 4.3%; Score 65; DB 4; Length 200;
Best Local Similarity 65.5%; Pred. No. 4.2e-11;
Matches 95; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 5 GATCTAGAACAGCAAGTCAAGTCAAGTCAATCTCTCACTCAGTGGTGGTGGTGGTGGT 64
Db 56 GATCTTGAAGCTGAACAGGTGAAGTAAATTCCTTAATCTCACTCAGTGGTGGTGGTGGT 115
QY 65 GAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGCAAACTTAAGTATTGGGAGAT 124
Db 116 GAACACAGTGGTGGAGAGCGCTTACAGCTATCTTAGAAGACCGATTACAGAACTGGGTGAG 175
QY 125 CGATGGGCAAACTCTGTAGATGGA 149
Db 176 CGTGGACAGCAGTATGCCGTGGA 200

RESULT 11

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
; US-08-232-463-14

Query Match 3.6%; Score 53.4; DB 1; Length 7218;
Best Local Similarity 3.8%; Pred. No. 2.8e-06;
Matches 15; Conservative 220; Mismatches 156; Indels 0; Gaps 0;
QY 242 AAAGATGATGAATTAAAGCGGCGACCTATTGGAGCGACTTCCAGCAGTTTCAGAA 301
DB 1436 ACRR 1377
QY 302 CAGAACGATGATACATAGGCGCTTCAAGAGGGGAATTGAAACTAAGAACCTGTATCATG 361
DB 1376 RRR 1317
QY 362 AGTACTCTTGACACTGACGAATATTTCTGACAGAGCAGCCTTTGGAGGAGCTAGAGAA 421
DB 1316 RRR 1257
QY 422 CTCTACCGAGGCCCGCAGAGCTGCTCTCTGAGAGAGAGCCAGAAATGTCACCTCGGCTT 481
DB 1256 RRR 1197
QY 482 CTACGAACAGCGCTGAGAGAGTCAATAGTGTGGGAAATGAACTGCACCTCCGCT 541

DB 1196 RRR 1137
QY 542 GACTGCGACAGAAATAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGCCAG 601
DB 1136 RRR 1077
QY 602 GATGAGCTGGACCTCAAGCTGCGCAAGCTG 632
DB 1076 RRRRRRRRRATCGCAAGCTCCCTCGACCTG 1046

RESULT 12

US-09-668-313A-10
; Sequence 10, Application US/09668313A
; Patent No. 6503756
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRESSION
; FILE REFERENCE: RTS-0127
; CURRENT APPLICATION NUMBER: US/09/668,313A
; CURRENT FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 247
; SEQ ID NO 10
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (218)...(1891)
; US-09-668-313A-10

Query Match 2.9%; Score 44.2; DB 4; Length 2574;
Best Local Similarity 54.7%; Pred. No. 0.0014;
Matches 88; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 953 TGGGAGAGAGCCATCTCGCCAAACAAAGTCCCTACTATATCAACACGAGACTCAAAACA 1012
DB 1733 TGGGAGAGAGCTTACACAGCAGATGGAAATCAAGTACTTCTCATCAACACGAGACTGACAGACC 1792
QY 1013 ACTTGTCTGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGCTCACTGGAATAAT 1072
DB 1793 ACCTCTCTGGATCCACCCCGTGATGAGCGCCCTGAACTCTCTGTGTCAGAGGAGAGTGAA 1852
QY 1073 GTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAG 1113
DB 1853 GAGGACTGTCCAGAGAGCTAAACAGACCCCGAAAAGCTGATG 1893

RESULT 13

US-09-620-312D-69
; Sequence 69, Application US/09620312D
; Patent No. 6596662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Felyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yuning
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 659662el Nucleic Acids and

```
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 69
; LENGTH: 1690
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128)..(1522)
; US-09-620-312D-69

Query Match      2.9%; Score 42.8; DB 4; Length 1690;
Best Local Similarity 47.9%; Pred. No. 0.003;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;

QY 580 CCAGAACTTCAAGAGCCACCGATGAGCTGGACCTCAAGTGGCCGCAAGCTGAGTGCAT 639
Db      |||||
QY 55 CAAGGAGTTGACCAAGTGGCGCACGACCTGGACGACGAGCTGGCATGGGTTCAGGAGCG 114
Db      |||||
QY 640 CAAGGGA---TCCTGGACCGCGTGGGGATCTCTCATTTGACTCTCTCCAGATCACCT 696
Db      |||||
QY 115 GCTGCCATGGCCATGACAGAGGAGGCAACGGTTTGAGGCGGTTCAGCAGCACAT 174
Db      |||||
QY 697 CGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAAACGTGAGCCACGT 756
Db      |||||
QY 175 CAAGAAAGAACAGGCGCTGGCGGAGATCCAGCGCATGGCGCGCGCTGGAGGAGGT 234
Db      |||||
QY 757 CAATGACCTTCTGCGCAGCTTACACTTTGGGCATTCAGCTCTCAGCGTATACCTCAG 816
Db      |||||
QY 235 GCTGAGCGCGCGCGCGCTGGCGTTCGTCGCGAGCGCGCGGAGGAGGAGGAGTGCACG 294
Db      |||||
QY 817 CACTCTGGAAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCGCTGAGGAGCGAGT 876
Db      |||||
QY 295 GGGCTTGAGCAGCTGACAGCGCGCTGGCGGAGTCCAGCGAGCTCGGAGGCTCCGAGCGGCA 354
Db      |||||
QY 877 CAGGAGCTGATGAAGCCACAGG 902
Db      |||||
QY 355 CGAGGTGCTGAGCGCGCTTCAGG 380
Db      |||||

RESULT 14
US-09-368-590-1
; Sequence 1, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; CURRENT FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7812
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(6879)
; NAME/KEY: unsure
; LOCATION: (100)...(102)
; NAME/KEY: unsure

; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 69
; LENGTH: 1690
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128)..(1522)
; US-09-620-312D-69

Query Match      2.9%; Score 42.8; DB 3; Length 7812;
Best Local Similarity 47.9%; Pred. No. 0.0083;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;

QY 580 CCAGGAACTTCAAGAGCCACCGATGAGCTGGACCTCAAGTGGCCGCAAGCTGAGTGCAT 639
Db      |||||
QY 3639 CAAGGAGTTGACCAAGTGGCGCACGCTGACGAGCTGGCATGGGTTCAGGAGCG 3698
Db      |||||
QY 640 CAAGGGA---TCCTGGACCGCGTGGGGATCTCTCATTTGACTCTCTCCAGATCACCT 696
Db      |||||
QY 3699 GCTGCCATGGCCATGACAGAGGAGGCAACGGTTTGAGGCGGTTCAGGAGCGCAT 3758
Db      |||||
QY 697 CGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAAACGTGAGCCACGT 756
Db      |||||
QY 3759 CAAGAAAGAACAGGCGCTGGCGGAGATCCAGCGCATGGCGCGCGCTGGAGGAGGT 3818
Db      |||||
QY 757 CAATGACCTTCTGCGCAGCTTACACTTTGGGCATTCAGCTCTCAGCGTATACCTCAG 816
Db      |||||
QY 3819 GCTGAGCGCGCGCGCGCTGGCGTTCGTCGCGAGCGCGGAGGAGGAGGAGTGCACG 3878
Db      |||||
QY 817 CACTCTGGAAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCGCTGAGGAGCGAGT 876
Db      |||||
QY 3879 GGGCTTGAGCAGCTGACAGCGCGCTGGCGGAGTCCAGCGCATGGCGCGCGCTGGAGGAGT 3938
Db      |||||
QY 877 CAGGAGCTGATGAAGCCACAGG 902
Db      |||||
QY 3939 GCAGGTGCTGAGCGCGCTTCAGG 3964
Db      |||||

RESULT 15
US-08-425-069-3
; Sequence 3, Application US/08425069
; Patent No. 5728810
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V.
; APPLICANT: Xu, Ming
; APPLICANT: Hinman, Michael B.
; TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
; TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
; TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 No. 5728810th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/425,069
; FILING DATE: 19-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1447-106P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1995 base pairs
```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; IMMEDIATE SOURCE:
; CLONE: p68
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1785
; US-08-425-069-3

Query Match 2.6%; Score 38.6; DB 1; Length 1995;
Best Local Similarity 44.0%; Pred. No. 0.078;
Matches 164; Conservative 0; Mismatches 209; Indels 0; Gaps 0;
QY 290 GCAGTTTCAGAGCAGACGATGTACATAGGGCCCTTCAAGAGGGAATTGAACTAAGAA 349
Db |||||
731 GCAGTGCAGCTGCAGCGCCGACGAGACCTGGACAACAAAGGACCCGGAGGATATGGAC 790
QY 350 CCTGTAATCATGACTCTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAA 409
Db |||||
791 CAGGACAAACAGGACCATCTGGACCCGGTAGTGCCTGAGCAGCAGCCGCCGACGAG 850
QY 410 GGACTAGAGAACTCTTACCAGGAGCCGACAGAGCTGCTCTCTGAGGAGAGCCCAAGAT 459
Db |||||
851 GACCTGGAGGATATGGCCCTGGACACAGAGGACCCGAGGATATGGACCAGGACAAACAAG 910
QY 470 GTCACTCGGCTTTACGAAGCAGGCTGAGGAGTCAATCTGAGTGGGAAAATTGAAC 529
Db |||||
911 GACCATCTGGAGCAGGCGAGTGCAGCAGCAGCGCCGACGAGACCTGGACACACAGGAT 970
QY 530 CTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTT 589
Db |||||
971 TAGGAGTTATGGACCAAGGACCAACAGGTCCAGGAGGATATGGACCAGGACAAACAGGTC 1030
QY 590 CAAGAGCCCGATGAGCTGGACCTCAAGCTGCCCAAGCTGAGGTGATCAAGGGATCC 649
Db |||||
1031 CAGGAGGATATGGACCAAGGATGCTCTCTGACGAGCAGCGCCGACGAGGACCTGGACAC 1090
QY 650 TGGCAGCCCGTGG 662
Db |||||
1091 AAGGACCAGGAGG 1103

Search completed: September 19, 2004, 11:06:59
Job time : 81.4052 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 19, 2004, 08:30:27 ; Search time 500.133 Seconds
(without alignments)
15148.395 Million cell updates/sec

Title: US-09-845-416-12_COPY_1500_3000

Perfect score: 1501

Sequence: 1 agaagatctagaacaagac.....ttctgcatgattctatccaa 1501

Scoring table: IDENTITY_NUC
Gapex 10.0 , Gapext 1.0

Searched: 3327077 seqs, 2523723180 residues

Total number of hits satisfying chosen parameters: 6654154

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*

15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*

16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*

17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1501	100.0	3510	10	US-09-845-416-12
2	1501	100.0	4476	10	US-09-845-416-31
3	1328.8	88.5	2169	10	US-09-845-416-4
4	1328.8	88.5	3531	10	US-09-845-416-10
5	1328.8	88.5	3858	10	US-09-845-416-9
6	1328.8	88.5	3999	10	US-09-845-416-6
7	1328.8	88.5	4182	10	US-09-845-416-2
8	1328.8	88.5	4498	10	US-09-845-416-30
9	1328.8	88.5	4825	10	US-09-845-416-29
10	1328.8	88.5	4848	10	US-09-845-416-35
11	1328.8	88.5	4966	10	US-09-845-416-28
12	1328.8	88.5	4990	10	US-09-845-416-34
13	1328.8	88.5	5060	10	US-09-845-416-36
14	1328.8	88.5	5149	10	US-09-845-416-27

Sequence 40, Appl
Sequence 41, Appl
Sequence 42, Appl
Sequence 1, Appl
Sequence 44, Appl
Sequence 47, Appl
Sequence 22, Appl
Sequence 2284, Ap
Sequence 1, Appl
Sequence 434, App
Sequence 434, App
Sequence 981, App
Sequence 981, App
Sequence 108, App
Sequence 13, Appl
Sequence 2, Appl
Sequence 14, Appl
Sequence 32, Appl
Sequence 39, Appl
Sequence 15, Appl
Sequence 1598, Ap
Sequence 4, Appl
Sequence 23, Appl
Sequence 3, Appl
Sequence 667, App
Sequence 157, App
Sequence 35, Appl
Sequence 32, Appl
Sequence 33, Appl
Sequence 34, Appl
Sequence 3, Appl

US-10-149-736-40
US-10-149-736-41
US-10-149-736-42
US-09-845-416-1
US-10-149-736-44
US-10-149-736-47
US-09-782-378A-22
US-09-880-107-2284
US-10-149-736-1
US-10-149-736-1
US-10-342-887-434
US-10-172-118-434
US-10-342-887-981
US-10-172-118-981
US-10-341-434-108
US-10-845-416-13
US-10-149-736-2
US-09-845-416-14
US-09-845-416-32
US-10-149-736-39
US-09-845-416-15
US-10-152-319A-1598
US-10-149-736-4
US-09-782-378A-23
US-10-149-736-3
US-10-101-510-667
US-10-220-120-157
US-10-149-736-35
US-10-149-736-32
US-10-149-736-33
US-10-149-736-34
US-09-845-416-3

15 1328.8 88.5 5339 16
16 1328.8 88.5 5462 16
17 1328.8 88.5 8689 16
18 1328.8 88.5 11058 10
19 1328.8 88.5 11443 16
20 1328.8 88.5 12057 16
21 1328.8 88.5 13957 9
22 1328.8 88.5 13957 9
23 1328.8 88.5 13957 16
24 1328.8 88.5 14069 13
25 1328.8 88.5 14069 13
26 1328.8 88.5 14082 13
27 1328.8 88.5 14082 13
28 1328.8 88.5 14082 16
29 1324 88.2 1821 10
30 1149.6 76.6 13815 16
31 1041 69.4 3446 10
32 1041 69.4 4414 10
33 1011.6 67.4 5417 16
34 937 62.4 1434 10
35 555.6 37.0 10705 12
36 554.2 36.9 11096 16
37 547.4 36.5 10302 9
38 547.4 36.5 10302 16
39 538.6 35.9 16531 15
40 527.4 35.1 5106 13
41 393 26.2 887 16
42 387 25.8 387 16
43 324 21.6 224 16
44 216 14.4 216 16
45 185.4 12.4 1991 10

ALIGNMENTS

RESULT 1

US-09-845-416-12
; Sequence 12, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12
; LENGTH: 3510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-12

Query Match 100.0%; Score 1501; DB 10; Length 3510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGATCTAGAACAGACAGTACGGGTCAATTCCTCACTCACATGCTGGTGTACT 60
1500 AGAAGATCTAGAACAGACAGTACGGGTCAATTCCTCACTCACATGCTGGTGTACT 1559

QY 61 TGATGAATCTAGTGGAGATCAGCAACTGCTCTTTTGAAGAACAACTTAAGGTATTGG 120
1560 TGATGAATCTAGTGGAGATCAGCAACTGCTCTTTTGAAGAACAACTTAAGGTATTGG 1619

QY 121 AGATCGATGGCAACATCTGTAGTGGACAGACCGCTGGTCTTTTACAGACAG 180
1620 AGATCGATGGCAACATCTGTAGTGGACAGACCGCTGGTCTTTTACAGACAG 1679

421 ACTCTACGAGCCGACAGAGCTGCTCTCTGAGGAGAGCCAGATGTCACCTGGCT 480
Db ACTCTACGAGCCGACAGAGCTGCTCTCTGAGGAGAGCCAGATGTCACCTGGCT 2735
481 TCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGC 540
Db TCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGC 2795
541 TGACTGCGAGAAAAATAGATGAGAGCCCTTGAAGACTCAGGAATTCAGAGGCCAC 600
Db TGACTGCGAGAAAAATAGATGAGAGCCCTTGAAGACTCAGGAATTCAGAGGCCAC 2855
601 GGATGAGTGCACCTCAAGCTGCCCAAGCTGAGTGCATCAAGGGATCTCTGGAGCCCGT 660
Db GGATGAGTGCACCTCAAGCTGCCCAAGCTGAGTGCATCAAGGGATCTCTGGAGCCCGT 2915
661 GGGCGATCTCTCAATGACTCTCTCCAAAGTACCTCGAGAAAGTCAAGGCACTTCGAGG 720
Db GGGCGATCTCTCAATGACTCTCTCCAAAGTACCTCGAGAAAGTCAAGGCACTTCGAGG 2975
721 AGAATTCGCGCTCTGAAGAGGACGTGAGCCAGCTCAATGACCTTCTCGCAGCTTAC 780
Db AGAATTCGCGCTCTGAAGAGGACGTGAGCCAGCTCAATGACCTTCTCGCAGCTTAC 3035
781 CACTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACCCAG 840
Db CACTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACCCAG 3095
841 ATGGAAGCTTTCGAGTGGCCGCTGAGGACCGAGTCAGGAGCTGCATGAAGCCACAG 900
Db ATGGAAGCTTTCGAGTGGCCGCTGAGGACCGAGTCAGGAGCTGCATGAAGCCACAG 3155
901 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTCTGTCAGGGTCCCTGGGAG 960
Db GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTCTGTCAGGGTCCCTGGGAG 3215
961 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACAGAGCTCAAACTTCGCTG 1020
Db AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACAGAGCTCAAACTTCGCTG 3275
1021 GGACATCCCAAAATGACAGAGCTCTACAGCTCTTTAGCTGACCTGAATATGTCAGATT 1080
Db GGACATCCCAAAATGACAGAGCTCTACAGCTCTTTAGCTGACCTGAATATGTCAGATT 3335
1081 CTCAGCTTATAGGACTGCCATGAATCCGAGACTGCGAGAGCCCTTTGCTTGGATCT 1140
Db CTCAGCTTATAGGACTGCCATGAATCCGAGACTGCGAGAGCCCTTTGCTTGGATCT 3395
1141 CTTGAGCCTGTGAGTGTGATGCTTGGACGAGCAACCTCAGCAAAATGACCA 1200
Db CTTGAGCCTGTGAGTGTGATGCTTGGACGAGCAACCTCAGCAAAATGACCA 3455
1201 GCGCATGGATATCTCTGAGATTAATTTGTTGACCACTATTTATGACCGCTGGAGCA 1260
Db GCGCATGGATATCTCTGAGATTAATTTGTTGACCACTATTTATGACCGCTGGAGCA 3515
1261 AGAGCAACAATTTGGTCAAGTCCCTCTCTGCGTGATATGCTGAACTGGCTGCT 1320
Db AGAGCAACAATTTGGTCAAGTCCCTCTCTGCGTGATATGCTGAACTGGCTGCT 3575
1321 GAATGTTTATGATACGGACGAAACAGGAGGATCCGTGCTGCTTTTAAACTGGCAT 1380
Db GAATGTTTATGATACGGACGAAACAGGAGGATCCGTGCTGCTTTTAAACTGGCAT 3635
1381 CATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 1440
Db CATTTCCCTGTGTAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 3695
1441 AAGTTCAACAGATTTTGTACCGAGCGAGCTGGGCTCTCTGCAATGATTTATCCA 1500
Db AAGTTCAACAGATTTTGTACCGAGCGAGCTGGGCTCTCTGCAATGATTTATCCA 3755
1501 A 1501

Db 3756 A 3756

RESULT 3

US-09-845-416-4
; Sequence 4, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-4

Query Match 88.5%; Score 1328.8; DB 10; Length 2169;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 162 GGGTCTTTTACAAGACAGTCTGACCCAGTGAAGCGTCTGCACCTTCTCTGCAGGAAC 221
Db 333 GGTCCCATTTGGAAGCCAGTCTGACCCAGTGAAGCGTCTGCACCTTCTCTGCAGGAAC 392
QY 222 TTCTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGCGGCGAGGCCCTTCAAGAGGAAATTGAAA 281
Db 393 TTCTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGCGGCGAGGCCCTTCAAGAGGAAATTGAAA 452
QY 282 ACTTCCAGCAGTTGAAAGCAGACGATGATACATAGGSCCTTCAAGAGGAAATTGAAA 341
Db 453 ACTTCCAGCAGTTGAAAGCAGACGATGATACATAGGSCCTTCAAGAGGAAATTGAAA 512
QY 342 CTAAGAAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 401
Db 513 CTAAGAAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 572
QY 402 CTTTGAAGAGACTAGAGAAACTCTACAGAGGCCAGAGAGCTGCTCTCTGAGGAGAGAG 461
Db 573 CTTTGAAGAGACTAGAGAAACTCTACAGAGGCCAGAGAGAGCTGCTCTCTGAGGAGAGAG 632
QY 462 CCCAGATCTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTGAGTGGAAA 521
Db 633 CCCAGATCTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTGAGTGGAAA 592
QY 522 AATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCC 581
Db 693 AATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCC 752
QY 582 AGGAATCTCAAGAGGCCAGGATGAGCTGGAACCTCAAGCTGGCCAGCTGAGGTGATCA 641
Db 753 AGGAATCTCAAGAGGCCAGGATGAGCTGGAACCTCAAGCTGGCCAGCTGAGGTGATCA 812
QY 642 AGGGATCTTGGCAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCTCAAGATCACTTCGAGA 701
Db 813 AGGGATCTTGGCAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCTCAAGATCACTTCGAGA 872
QY 702 AAGTCAAGGCACTTCGAGGAGAAATTGCGCTCTGAAAGAGAACCTGAGGCCAGCTCAATG 761
Db 873 AAGTCAAGGCACTTCGAGGAGAAATTGCGCTCTGAAAGAGAACCTGAGGCCAGCTCAATG 932
QY 762 ACCTTCTCGCCAGCTTACCACTTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTC 821
Db 933 ACCTTCTCGCCAGCTTACCACTTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTC 992

QY 822 TGGAGAGCTGAACACACAGATGGAAGCTTCTGACGCTGCGCTCGAGACCGAGTCAAGC 881
DB 993 TGGAGAGCTGNACACAGATGGAAGCTTCTGACGCTGCGCTCGAGACCGAGTCAAGC 1052
QY 882 AGCTGCATGAAGCCACAGGACATTTGCTCAGCATCTCAGCATTTCTTTCCAGCTG 941
DB 1053 AGCTGCATGAAGCCACAGGACATTTGCTCAGCATCTCAGCATTTCTTTCCAGCTG 1112
QY 942 TCCAGGGTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACG 1001
DB 1113 TCCAGGGTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACG 1172
QY 1002 AGACTCAAAACACTTGTGGGACCATCCCAAAATGACAGACTCTACAGTCTTTAGCTG 1061
DB 1173 AGACTCAAAACACTTGTGGGACCATCCCAAAATGACAGACTCTACAGTCTTTAGCTG 1232
QY 1062 ACCTGAATAATGTAGATTTCTAGCTTATAGGACTGCCATGAAACTCGGAAGACTGCAGA 1121
DB 1233 ACCTGAATAATGTAGATTTCTAGCTTATAGGACTGCCATGAAACTCGGAAGACTGCAGA 1292
QY 1122 AGGCCCTTTGCTTGGATCTCTTGGCCCTGTGAGCTGTGATGTGATGCTTGGACAGCACA 1181
DB 1293 AGGCCCTTTGCTTGGATCTCTTGGCCCTGTGAGCTGTGATGTGATGCTTGGACAGCACA 1352
QY 1182 ACCTCAAGAAAATGACAGGCCATGATATCTCTGAGATTAATTAATTTGACCACTA 1241
DB 1353 ACCTCAAGAAAATGACAGGCCATGATATCTCTGAGATTAATTAATTTGACCACTA 1412
QY 1242 TTTATGACCGCTGGAGCAAGACACACATTTGGTCAAGCTCCCTCTCTGGGTGGATA 1301
DB 1413 TTTATGACCGCTGGAGCAAGACACACATTTGGTCAAGCTCCCTCTCTGGGTGGATA 1472
QY 1302 TGTGCTGAACCTGGCTGTGAATTTTATGATACGGGACGAAACAGGAGGATCCGTGTCC 1361
DB 1473 TGTGCTGAACCTGGCTGTGAATTTTATGATACGGGACGAAACAGGAGGATCCGTGTCC 1532
QY 1362 TGTCTTTAAACTGGCATATTTCCCTGTGTAAGACACATTTGGAAGACAACTACAGAT 1421
DB 1533 TGTCTTTAAACTGGCATATTTCCCTGTGTAAGACACATTTGGAAGACAACTACAGAT 1592
QY 1422 ACCTTTTCAGCAAGTGGCAAGTTTCAACAGGATTTTGTACCGAGCGAGGCTGGGCTCC 1481
DB 1593 ACCTTTTCAGCAAGTGGCAAGTTTCAACAGGATTTTGTACCGAGCGAGGCTGGGCTCC 1652
QY 1482 TCTGCATGATCTATCCAA 1501
DB 1653 TCTGCATGATCTATCCAA 1672

RESULT 4

US-09-845-416-10
; Sequence 10, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 3531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-10

Query Match 88.5%; Score 1328.8; DB 10; Length 3531;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 162 GGGTTCTTTTAAAGACAGTTCTGACCACTGGAAGCTTCTGACCTTCTCTGACGAGAAC 221
DB 1682 GGTCCCATTTTGAAGCCAGTTCTGACCACTGGAAGCTTCTGACCTTCTCTGACGAGAAC 1741
QY 222 TTCTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGGACACTTATTGAGGCG 281
DB 1742 TTCTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGGACACTTATTGAGGCG 1801
QY 282 ACTTTCCAGCACTGAGAAAGCAAGATGTACATAGGCTTTCAAGAGGGAATTTGAAAA 341
DB 1802 ACTTTCCAGCACTGAGAAAGCAAGATGTACATAGGCTTTCAAGAGGGAATTTGAAAA 1861
QY 342 CTAAAGAACTCTAATCATGAGTACTCTTTGAGACTGTAGCAATATTTCTGACAGAGCAGC 401
DB 1862 CTAAAGAACTCTAATCATGAGTACTCTTTGAGACTGTAGCAATATTTCTGACAGAGCAGC 1921
QY 402 CTTTGAAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCTCTCTGAGGAGAGAG 461
DB 1922 CTTTGAAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCTCTCTGAGGAGAGAG 1981
QY 462 CCCAGATGTCACTCGGCTTTCTACGAAGCAGGCTGAGAGGTCAATACTGAGTGGGAAA 521
DB 1982 CCCAGATGTCACTCGGCTTTCTACGAAGCAGGCTGAGAGGTCAATACTGAGTGGGAAA 2041
QY 522 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCC 581
DB 2042 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCC 2101
QY 582 AGGAACCTCAAGAGGCCAGGATGAGCTGAGACTCAAGCTGGGCCAAGCTGAGTGTATCA 641
DB 2102 AGGAACCTCAAGAGGCCAGGATGAGCTGAGACTCAAGCTGGGCCAAGCTGAGTGTATCA 2161
QY 642 AGGGATCTCGGAGCCCGTGGCGCATCTCTCATTTGACTCTCTCAAGATCACTCCGAGA 701
DB 2162 AGGGATCTCGGAGCCCGTGGCGCATCTCTCATTTGACTCTCTCAAGATCACTCCGAGA 2221
QY 702 AAGTCAAGGCACTTCGAGGAGAAATTTGGCTCTGTAAGAGAAAGCTGAGCCACGTCATG 761
DB 2222 AAGTCAAGGCACTTCGAGGAGAAATTTGGCTCTGTAAGAGAAAGCTGAGCCACGTCATG 2281
QY 762 ACCTTGTCTGCCAGCTTACCACTTTGGCATTTCAGCTCTCAAGCTGATATCACTCAGCCTC 821
DB 2282 ACCTTGTCTGCCAGCTTACCACTTTGGCATTTCAGCTCTCAAGCTGATATCACTCAGCCTC 2341
QY 822 TGGAGAGCTGAACACACAGATGGAAGCTTCTGACGCTGCGCTCGAGGACCGAGTCAAGC 881
DB 2342 TGGAGAGCTGAACACACAGATGGAAGCTTCTGACGCTGCGCTCGAGGACCGAGTCAAGC 2401
QY 882 AGCTGCATGAGCCACAGGAGACTTTGGTCCAGCATCTCAGCAGCTTTCTTCCAGCTG 941
DB 2402 AGCTGCATGAGCCACAGGAGACTTTGGTCCAGCATCTCAGCAGCTTTCTTCCAGCTG 2461
QY 942 TCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACACG 1001
DB 2462 TCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACACG 2521
QY 1002 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGACTCTACAGCTTTAGCTG 1061
DB 2522 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGACTCTACAGCTTTAGCTG 2581
QY 1062 ACCTGAATAATGTAGATTTCTCAGCTTATAGGCTGCCATGAAACTCGGAAGACTGCAGA 1121
DB 2582 ACCTGAATAATGTAGATTTCTCAGCTTATAGGCTGCCATGAAACTCGGAAGACTGCAGA 2641
QY 1122 AGGCCCTTTGCTTGGATCTCTTGGCCCTGTGAGCTGTGATGTGATGCTTGGACAGCACA 1181
DB 2642 AGGCCCTTTGCTTGGATCTCTTGGCCCTGTGAGCTGTGATGTGATGCTTGGACAGCACA 2701
QY 1182 ACCTCAAGAAAATGACAGGCCATGATATCTCTGAGATTAATTAATTTGTTGACCACTA 1241
DB 2702 ACCTCAAGAAAATGACAGGCCATGATATCTCTGAGATTAATTAATTTGTTGACCACTA 2761


```
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3999
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-6

Query Match      88.5%; Score 1328.8; DB 10; Length 3999;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 162 GGGTTCTTTTACAAGACAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAGC 221
Db 2150 GGTCCCAATTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAGC 2209
QY 222 TTCTGGTGTGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGGACCTATTGGAGCG 281
Db 2210 TTCTGGTGTGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGGACCTATTGGAGCG 2269
QY 282 ACTTTCCAGCAGTTTCAAGACAGAACGATGACATAGGGCTTCAAGAGGGAATTGAAAA 341
Db 2270 ACTTTCCAGCAGTTTCAAGACAGAACGATGACATAGGGCTTCAAGAGGGAATTGAAAA 2329
QY 342 CTAAGAACCCTGTAATCATAGTACTCTTGAGACTGACGAATATTTCTGACAGAGCAGC 401
Db 2330 CTAAGAACCCTGTAATCATAGTACTCTTGAGACTGACGAATATTTCTGACAGAGCAGC 2389
QY 402 CTTTGGAGGAGTACAGAACTCTACCAAGAGCCAGAGAGTGCCTCTGAGGAGAG 461
Db 2390 CTTTGGAGGAGTACAGAACTCTACCAAGAGCCAGAGAGTGCCTCTGAGGAGAG 2449
QY 462 CCCAGAACTGTCACCTGGGTTCTACGAAGCAGGCTGAGGAGGTCAATCTAGTGGGAAA 521
Db 2450 CCCAGAACTGTCACCTGGGTTCTACGAAGCAGGCTGAGGAGGTCAATCTAGTGGGAAA 2509
QY 522 AATTGAACCTGCACCTCCGCTGACCTGGCAGAGAAAATAGATGAGACCTTTGAAGACTCC 581
Db 2510 AATTGAACCTGCACCTCCGCTGACCTGGCAGAGAAAATAGATGAGACCTTTGAAGACTCC 2569
QY 582 AGGAACCTTCAAGAGGCCACCGATGAGTGGACCTCAAGCTGCGCAAGCTGAGGTGATCA 641
Db 2570 AGGAACCTTCAAGAGGCCACCGATGAGTGGACCTCAAGCTGCGCAAGCTGAGGTGATCA 2629
QY 642 AGGATCTTGGCAGCCGCTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACTGAGAGA 701
Db 2630 AGGATCTTGGCAGCCGCTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACTGAGAGA 2689
QY 702 AAGTCAAGCAGCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTCGAGCCAGTCAATG 761
Db 2690 AAGTCAAGCAGCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTCGAGCCAGTCAATG 2749
QY 762 ACCTTGTCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCCTATTAACCTCAGCACTC 821
Db 2750 ACCTTGTCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCCTATTAACCTCAGCACTC 2809
QY 822 TGAAGACCTCAACACAGATGGAAGCTTCTGCAAGTGGCGCTGAGGACCGAGTCAAGC 881
Db 2810 TGAAGACCTCAACACAGATGGAAGCTTCTGCAAGTGGCGCTGAGGACCGAGTCAAGC 2869
QY 882 AGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGCTG 941
Db 2870 AGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGCTG 2929
QY 942 TCCAGGTCCTCTGGAGAGAGCCTATCTGCCCAACAAAGTGCCTACTATATCAACACG 1001
Db 2930 TCCAGGTCCTCTGGAGAGAGCCTATCTGCCCAACAAAGTGCCTACTATATCAACACG 2989

1002 AGACTCAAAACAACTTCTGGGACCAATCCCAAAATACAGAGCTCTTACCAGTCTTTAGCTG 1061
2990 AGACTCAAAACAACTTCTGGGACCAATCCCAAAATACAGAGCTCTTACCAGTCTTTAGCTG 3049
1062 ACCTGAATAAATGTCAAGATTTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCAGA 1121
3050 ACCTGAATAAATGTCAAGATTTCTCAGCTTATAGACTGCCATGAAACTCCGAAGACTGCAGA 3109
1122 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGATGATGATGATGATGATGATGAT 1181
3110 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGATGATGATGATGATGATGATGATGAT 3169
1182 ACCTCAAGCAAAATGACCAAGCCCATGGATATCTTCAGATATTAAATTTGTTTGACCACTA 1241
3170 ACCTCAAGCAAAATGACCAAGCCCATGGATATCTTCAGATATTAAATTTGTTTGACCACTA 3229
1242 TTTATGACCGCTCGAGCAAGACGACAAATTTGGTCAACGTCCTCTCTGCGTGGATA 1301
3230 TTTATGACCGCTCGAGCAAGACGACAAATTTGGTCAACGTCCTCTCTGCGTGGATA 3289
1302 TGTGCTGTAACCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1361
3290 TGTGCTGTAACCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3349
1362 TGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT 1421
3350 TGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT 3409
1422 ACCTTTCAAGCAAGTGCAGAGTTCAACAGGATTTTGTGACCGGAGGCTGGGCTTCC 1481
3410 ACCTTTCAAGCAAGTGCAGAGTTCAACAGGATTTTGTGACCGGAGGCTGGGCTTCC 3469
1482 TTCTGCATGATTCTATCCAA 1501
3470 TTCTGCATGATTCTATCCAA 3489

RESULT 7
US-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-2

Query Match      88.5%; Score 1328.8; DB 10; Length 4182;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 162 GGGTTCTTTTACAAGACAGTTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAGC 221
Db 2333 GGTCCCAATTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAGC 2392
QY 222 TTCTGGTGTGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGGACCTATTGGAGCG 281
Db 2393 TTCTGGTGTGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGGACCTATTGGAGCG 2452
QY 282 ACTTTCCAGCAGTTTCAAGACAGAACGATGATGATGATGATGATGATGATGATGATGATGAT 341
```

Db 2453 ACTTCCAGAGTTCAAGACAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAA 2512
Qy 342 CTAAGAACCTGTAATCATGACTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGC 401
Db 2513 CTAAGAACCTGTAATCATGACTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGC 2572
Qy 402 CTTTGGAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCTCTCGAGAGAGAG 461
Db 2573 CTTTGGAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCTCTCGAGAGAGAG 2632
Qy 462 CCCAGAAATGCTACGCGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 521
Db 2633 CCCAGAAATGCTACGCGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 2692
Qy 522 AATTGAACCTGCACTCCGCTGACTGCGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCC 581
Db 2693 AATTGAACCTGCACTCCGCTGACTGCGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCC 2752
Qy 582 AGGAACCTCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGCGCAAGCTGAGGTGATCA 641
Db 2753 AGGAACCTCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGCGCAAGCTGAGGTGATCA 2812
Qy 642 AGGATCTGCGACCGCTGGCGCATCTCTCATTTGACTCTCTCCAAGATCACTCGAGA 701
Db 2813 AGGATCTGCGACCGCTGGCGCATCTCTCATTTGACTCTCTCCAAGATCACTCGAGA 2872
Qy 702 AAGTCAAGGCACTTCGAGGAGAAAAATGCGCTCTGAAAGAGAAAGCTGAGCCACGTCAATG 761
Db 2873 AAGTCAAGGCACTTCGAGGAGAAAAATGCGCTCTGAAAGAGAAAGCTGAGCCACGTCAATG 2932
Qy 762 ACCTTGCTGCGCAGCTTACCACTTTGGGCAATGAGCTCTCACCGTATAAAGCTGAGCACTC 821
Db 2933 ACCTTGCTGCGCAGCTTACCACTTTGGGCAATGAGCTCTCACCGTATAAAGCTGAGCACTC 2992
Qy 822 TGGAGACCTCAACACAGATGGAAGCTTCGACAGGTGGCGTGGAGGCGGAGTCAAGC 881
Db 2993 TGGAGACCTCAACACAGATGGAAGCTTCGACAGGTGGCGTGGAGGCGGAGTCAAGC 3052
Qy 882 AGCTGCATGAAGCCCAAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTG 941
Db 3053 AGCTGCATGAAGCCCAAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTG 3112
Qy 942 TCCAGGCTCCCTGGAGAGAGCCATCTGCCCAACAAAGTCCCTACTATATCAACCAAG 1001
Db 3113 TCCAGGCTCCCTGGAGAGAGCCATCTGCCCAACAAAGTCCCTACTATATCAACCAAG 3172
Qy 1002 AGACTCAAAACAACTTCTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTTAGCTG 1061
Db 3173 AGACTCAAAACAACTTCTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTTAGCTG 3232
Qy 1062 ACCTGAATAATGTCAAGATTCAGCTTATAGGACTGCGCATGAAACTCCGAAGACTCGAGA 1121
Db 3233 ACCTGAATAATGTCAAGATTCAGCTTATAGGACTGCGCATGAAACTCCGAAGACTCGAGA 3292
Qy 1122 AGGCCCTTTGCTTGATCTCTTGAGCCTGTCAGCTGCATGTGATGCTTTGGACCAAGACA 1181
Db 3293 AGGCCCTTTGCTTGATCTCTTGAGCCTGTCAGCTGCATGTGATGCTTTGGACCAAGACA 3352
Qy 1182 ACCTCAAGCAAAATGACAGCCCATGGATATCTGCAAGATTTAAATTTGTTGACCACTA 1241
Db 3353 ACCTCAAGCAAAATGACAGCCCATGGATATCTGCAAGATTTAAATTTGTTGACCACTA 3412
Qy 1242 TTTATGACCGCTGAGCAAGAGCACAAATTTGGTCAAGCTCCCTCTCTGCGTGGATA 1301
Db 3413 TTTATGACCGCTGAGCAAGAGCACAAATTTGGTCAAGCTCCCTCTCTGCGTGGATA 3472
Qy 1302 TGTGCTGAATGCTGCTGATGTTTATGATACGGGACGAACAGGAGGATCCGCTGTC 1361
Db 3473 TGTGCTGAATGCTGCTGATGTTTATGATACGGGACGAACAGGAGGATCCGCTGTC 3532
Qy 1362 TGTCTTTTAAACTGGGATCAATTCCTCTGTGTAAGCACTTTGGAAGCAAGTACAGAT 1421
Db 3533 TGTCTTTTAAACTGGGATCAATTCCTCTGTGTAAGCACTTTGGAAGCAAGTACAGAT 3592

Qy 1422 ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCTCC 1481
Db 3593 ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCTCC 3652
Qy 1482 TTCTGCATGATTTCTATCCAA 1501
Db 3653 TTCTGCATGATTTCTATCCAA 3672

RESULT 8

US-09-845-416-30
; Sequence 30, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 4498
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-845-416-30

Query Match 88.5%; Score 1328.8; DB 10; Length 4498;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 162 GGGTCTTTTACAAGACAGTCTGACCACTGGAAGCGTCTGCACCTTTCTGCGAGAAC 221
Db 2439 GGTCCATTGGAAGCCAGTTCTGACCACTGGAAGCGTCTGCACCTTTCTGCGAGAAC 2498
Qy 222 TTCTGTGTGGTACAGCTGGAAGATGATGAATTAAGCGGCGAGCACCATTGAGGCG 281
Db 2499 TTCTGTGTGGTACAGCTGGAAGATGATGAATTAAGCGGCGAGCACCATTGAGGCG 2558
Qy 282 ACTTTCACAGCTTCAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAA 341
Db 2559 ACTTTCACAGCTTCAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAA 2618
Qy 342 CTAAGAACTCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 401
Db 2619 CTAAGAACTCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 2678
Qy 402 CTTTGAAGAGCTAGAGAACTCTACCAAGAGCCAGAGAGCTGCTCTGAGGAGAGAG 461
Db 2679 CTTTGAAGAGCTAGAGAACTCTACCAAGAGCCAGAGAGCTGCTCTGAGGAGAGAG 2738
Qy 462 CCCAGAAATGCTACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAA 521
Db 2739 CCCAGAAATGCTACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAA 2798
Qy 522 AATTGAACCTGCACTCCGCTGACTGCGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCC 581
Db 2799 AATTGAACCTGCACTCCGCTGACTGCGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCC 2858
Qy 582 AGGAACCTCAAGAGGCCACGGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGGTGATCA 641
Db 2859 AGGAACCTCAAGAGGCCACGGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGGTGATCA 2918
Qy 642 AGGATCTGCGACCGCTGGCGCATCTCTCATTTGACTCTCTCCAAGATCACTCGAGA 701
Db 2919 AGGATCTGCGACCGCTGGCGCATCTCTCATTTGACTCTCTCCAAGATCACTCGAGA 2978
Qy 702 AAGTCAAGGCACTTCAGAGAGAAATTTGCCCTCTGAAAGAGCAAGTACAGTCAATG 761

Db 2979 AAGTCAAGGCACTTCAGAGAGAAATTGCGCTCTGAAAGAGAACGTCAGGCAACGTCATG 3038
QY 762 ACCTTGCTCGCAGCTTACCACTTTGGGCAATCAGCTCTACCGTATTAACCTCAGCACTC 821
Db 3039 ACCTTGCTCGCAGCTTACCACTTTGGGCAATCAGCTCTACCGTATTAACCTCAGCACTC 3098
QY 822 TGGAAAGACCTGAAACACCAAGATGGAAGCTTCTGCAAGTGCGCGTTCAGGACCGAGTCAGGC 881
Db 3099 TGGAAAGACCTGAAACACCAAGATGGAAGCTTCTGCAAGTGCGCGTTCAGGACCGAGTCAGGC 3158
QY 882 AGCTGCATGAAGCCACACAGGACCTTTGGTTCAGGATCTCAGCACTTTCTTTCCAGCTG 941
Db 3159 AGCTGCATGAAGCCACACAGGACCTTTGGTTCAGGATCTCAGCACTTTCTTTCCAGCTG 3218
QY 942 TCCAGGTCCTTCGGGAGAGACCATCTCGCCAAACAAAGTCCCTACTATATCAACCAAG 1001
Db 3219 TCCAGGTCCTTCGGGAGAGACCATCTCGCCAAACAAAGTCCCTACTATATCAACCAAG 3278
QY 1002 AGACTCAAAACAACTTGTGGGACCAATCCCAAAATGACAGAGCTCTACCACTTTTAGCTG 1061
Db 3279 AGACTCAAAACAACTTGTGGGACCAATCCCAAAATGACAGAGCTCTACCACTTTTAGCTG 3338
QY 1062 ACCTGAATAATGTCAGATCTCAGCTTATAGGAGTCCATGAAACTCCGAGAGTCGAGA 1121
Db 3339 ACCTGAATAATGTCAGATCTCAGCTTATAGGAGTCCATGAAACTCCGAGAGTCGAGA 3398
QY 1122 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCCCTTGGACCAACCA 1181
Db 3399 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCCCTTGGACCAACCA 3458
QY 1182 ACCTCAACCAAAATGACAGGACCAATGATATCTCTGAGATTAATATTTGTTGACCACTA 1241
Db 3459 ACCTCAACCAAAATGACAGGACCAATGATATCTCTGAGATTAATATTTGTTGACCACTA 3518
QY 1242 TTTATGACCGCTCGGACAGACCAACAAATTTGGTCAAGCTCCCTCTCTGCGTGATA 1301
Db 3519 TTTATGACCGCTCGGACAGACCAACAAATTTGGTCAAGCTCCCTCTCTGCGTGATA 3578
QY 1302 TGTGCTGAATGCTGCTGAATGTTATGATACGGGACCAACAGGAGGATCCGTGTCC 1361
Db 3579 TGTGCTGAATGCTGCTGAATGTTATGATACGGGACCAACAGGAGGATCCGTGTCC 3638
QY 1362 TGTCTTTAAACTGGGATCATTTCCCTGTGTAAGCACATTTGGAGACAGTACAGAT 1421
Db 3639 TGTCTTTAAACTGGGATCATTTCCCTGTGTAAGCACATTTGGAGACAGTACAGAT 3698
QY 1422 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGGACGAGCTGGGCTCC 1481
Db 3699 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGGACGAGCTGGGCTCC 3758
QY 1482 TTCTGCATGATTTCTATCCAA 1501
Db 3759 TTCTGCATGATTTCTATCCAA 3778

RESULT 9

US-09-845-416-29
; Sequence 29, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 4825
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-845-416-29
Query Match 88.5%; Score 1328.8; DB 10; Length 4825;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 162 GGGTCTTTTCAAGACAGCTTCTGACCAAGTGAAGGCTTCGACCTTTCTCTGAGGAAAC 221
Db 2766 GGTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGGCTTCGACCTTTCTCTGAGGAAAC 2825
QY 222 TTCTGGTGTGGCTCAGAGTGAAGATGATGAATTAAGCCGGCAGGACCTATTGAGGCG 281
Db 2826 TTCTGGTGTGGCTCAGAGTGAAGATGATGAATTAAGCCGGCAGGACCTATTGAGGCG 2885
QY 282 ACTTTCCAGAGCTCAGAGCAGAAAGATGATACATAGGGCTTTCAAGAGGGAATGAAAA 341
Db 2886 ACTTTCCAGAGCTCAGAGCAGAAAGATGATACATAGGGCTTTCAAGAGGGAATGAAAA 2945
QY 342 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 401
Db 2946 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 3005
QY 402 CTTTGGAGGACTGAGAAACTCTTACAGAGCCGACAGAGCTGCTCTCTGAGGAGAGAG 461
Db 3006 CTTTGGAGGACTGAGAAACTCTTACAGAGCCGACAGAGCTGCTCTCTGAGGAGAGAG 3065
QY 462 CCAGAAATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATACTCAGTGGGAAA 521
Db 3066 CCAGAAATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATACTCAGTGGGAAA 3125
QY 522 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCC 581
Db 3126 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCC 3185
QY 582 AGGAATCTTCAAGAGCCACGATGAGCTGAGCTCAAGCTCGGCCACAGCTGAGTGATCA 641
Db 3186 AGGAATCTTCAAGAGCCACGATGAGCTGAGCTCAAGCTCGGCCACAGCTGAGTGATCA 3245
QY 642 AGGGATCTTGGCAGCCCGCTGGCGCATCTCTCATTTGACTCTCTCAAGATCACCTCCAGA 701
Db 3246 AGGGATCTTGGCAGCCCGCTGGCGCATCTCTCATTTGACTCTCTCAAGATCACCTCCAGA 3305
QY 702 AAGTCAAGGCACTTGGAGGAAATTTGGCCTCTGAAAGAGAAAGTGGAGCCACGTCATG 761
Db 3306 AAGTCAAGGCACTTGGAGGAAATTTGGCCTCTGAAAGAGAAAGTGGAGCCACGTCATG 3365
QY 762 ACCTTGTCTGCGCAGCTTACCACTTTGGSCATTTAGCTCTCACCGTATTAACCTCAGCACTC 821
Db 3366 ACCTTGTCTGCGCAGCTTACCACTTTGGSCATTTAGCTCTCACCGTATTAACCTCAGCACTC 3425
QY 822 TGGAAAGACCTGAACACCAAGATGGAAGCTTTGCAAGTGGCGCTCGAGGACCGAGTCAGGC 881
Db 3426 TGGAAAGACCTGAACACCAAGATGGAAGCTTTGCAAGTGGCGCTCGAGGACCGAGTCAGGC 3485
QY 882 AGCTGCATGAAGCCACACAGGACCTTTGGTCCAGATCTCAGCACTTTCTTTCCAGCTG 941
Db 3486 AGCTGCATGAAGCCACACAGGACCTTTGGTCCAGATCTCAGCACTTTCTTTCCAGCTG 3545
QY 942 TCCAGGTCCTTCGGGAGAGACCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCAAG 1001
Db 3546 TCCAGGTCCTTCGGGAGAGACCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCAAG 3605
QY 1002 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGCTTTTAGCTG 1061
Db 3606 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGCTTTTAGCTG 3665
QY 1062 ACCTGAATAATGTCAGATTTCTCAGCTTATAGGATGCGCATGAAACTCCGAGAGTGCAGA 1121
Db 3666 ACCTGAATAATGTCAGATTTCTCAGCTTATAGGATGCGCATGAAACTCCGAGAGTGCAGA 3725
QY 1122 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGTCATGTGATGCCCTTGGACCAACCA 1181

Db 3726 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGACGTGATGATGCTTGGACGACACA 3785
Qy 1182 ACCTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATATTAATTTGTTGACCACTA 1241
Db 3786 ACCTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATATTAATTTGTTGACCACTA 3845
Qy 1242 TTTATGACCGCTGGAGAGAGACAAACAAATTTGGTCAACGTCCTCTCTGCGTGGATA 1301
Db 3846 TTTATGACCGCTGGAGAGAGACAAACAAATTTGGTCAACGTCCTCTCTGCGTGGATA 3905
Qy 1302 TGTCTCTGAATGCTGCTGAATCTTTATGATACGGGACGAAACAGGAGGATCCGTGTCC 1361
Db 3906 TGTCTCTGAATGCTGCTGAATCTTTATGATACGGGACGAAACAGGAGGATCCGTGTCC 3965
Qy 1362 TGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGAT 1421
Db 3966 TGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGAT 4025
Qy 1422 ACCTTTTCAGCAAGTGGCAAGTTTCACAGGATTTTGTGACCGGCGAGGCTGGGCTCC 1481
Db 4026 ACCTTTTCAGCAAGTGGCAAGTTTCACAGGATTTTGTGACCGGCGAGGCTGGGCTCC 4085
Qy 1482 TTCTGCATGATTCTATCCAA 1501
Db 4086 TTCTGCATGATTCTATCCAA 4105

RESULT 10
US-09-845-416-35
; Sequence 35, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: THEREOF
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 35
; LENGTH: 4848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-35

Query Match 88.5%; Score 1328.8; DB 10; Length 4848;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 162 GGGTCTTTTACAGACAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCTCTGAGGAAC 221
Db 2789 GGTCCCATTTGGAAGCCAGTTCTGACCAAGTGGAAAGCGTCTGCACCTTTCTCTGAGGAAC 2848
Qy 222 TTCTGCTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGGACCTATTTCGAGGCG 281
Db 2849 TTCTGCTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGGACCTATTTCGAGGCG 2908
Qy 282 ACTTTCCAGGAGTTCAGAAAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAA 341
Db 2909 ACTTTCCAGGAGTTCAGAAAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAA 2968
Qy 342 CTAAAGAACCTGTATCATGATGACTCTTTGAGACTGTACGATATTTTCGACAGCAGC 401
Db 2969 CTAAAGAACCTGTATCATGATGACTCTTTGAGACTGTACGATATTTTCGACAGCAGC 3028
Qy 402 CTTTTGAAGGACTAGAGAACTCTTACAGGAGCCAGAGAGCTGCTCTCTGAGGAGAG 461
Db 3029 CTTTGAAGGACTAGAGAACTCTTACAGGAGCCAGAGAGCTGCTCTCTGAGGAGAG 3088
Qy 462 CCCAGAAATGCTACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAA 521

RESULT 11

Db 3089 CCCAGAAATGCTACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAA 3148
Qy 522 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC 581
Db 3149 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC 3208
Qy 582 AGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCA 641
Db 3209 AGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCA 3268
Qy 642 AGGGATCTTGGCAGCCGCTGGGCGGATCTCTCATTTGACTCTCTCTCAAGATCACCTCGAGA 701
Db 3269 AGGGATCTTGGCAGCCGCTGGGCGGATCTCTCATTTGACTCTCTCTCAAGATCACCTCGAGA 3328
Qy 702 AAGTCAAGGCACCTTCAAGAGGAAATTTGGCCCTCTGAAAGAGAAACGTGAGCCACGTCATG 761
Db 3329 AAGTCAAGGCACCTTCAAGAGGAAATTTGGCCCTCTGAAAGAGAAACGTGAGCCACGTCATG 3388
Qy 762 ACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATTAACCTCAGCACTC 821
Db 3389 ACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATTAACCTCAGCACTC 3448
Qy 822 TGAAGAACCTGAACACCCAGATGGAAGCTTCTGACGCTGGCCCTCGAGGACCCAGTCAGGC 881
Db 3449 TGAAGAACCTGAACACCCAGATGGAAGCTTCTGACGCTGGCCCTCGAGGACCCAGTCAGGC 3508
Qy 882 AGCTGCATGAAGCCACAGGGAATTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTG 941
Db 3509 AGCTGCATGAAGCCACAGGGAATTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTG 3568
Qy 942 TCCAGGGTCCCTGGGAGAGAGCCATCTCCGCCAAAACAAAGTGCCTTACTATATCAACACG 1001
Db 3569 TCCAGGGTCCCTGGGAGAGAGCCATCTCCGCCAAAACAAAGTGCCTTACTATATCAACACG 3628
Qy 1002 AGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTG 1061
Db 3629 AGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAGCTG 3688
Qy 1062 ACCTGAATTAATGTGAGATTTCTCAGCTTATAGCACTGCCATGAAACTCCGAAAGCTGCAGA 1121
Db 3689 ACCTGAATTAATGTGAGATTTCTCAGCTTATAGCACTGCCATGAAACTCCGAAAGCTGCAGA 3748
Qy 1122 AGGCCCTTTGCTTGGATCTTTGAGCCTGTGACGTGTCATGTGATGCTTGGACCGACACA 1181
Db 3749 AGGCCCTTTGCTTGGATCTTTGAGCCTGTGACGTGTCATGTGATGCTTGGACCGACACA 3808
Qy 1182 ACCTGAAGCAAAATGACAGCCCATGGATATCTGCAGATTTAATTTGTTGACCACTA 1241
Db 3809 ACCTGAAGCAAAATGACAGCCCATGGATATCTGCAGATTTAATTTGTTGACCACTA 3868
Qy 1242 TTTATGACCGCTTGGAGAGAGACAAACAAATTTGGTCAACCTCTCTCTCGGTGGATA 1301
Db 3869 TTTATGACCGCTTGGAGAGAGACAAACAAATTTGGTCAACCTCTCTCTCGGTGGATA 3928
Qy 1302 TGTGCTGAACCTGGCTGTGAAATGTTTATGATACGGAGCAGAACAGGAGGATCCGTGTCC 1361
Db 3929 TGTGCTGAACCTGGCTGTGAAATGTTTATGATACGGAGCAGAACAGGAGGATCCGTGTCC 3988
Qy 1362 TGTCTTTTAAACTGSCATCATTTCCCTCTGTAAAGACACATTTGGAAGACAAAGTACAGAT 1421
Db 3989 TGTCTTTTAAACTGSCATCATTTCCCTCTGTAAAGACACATTTGGAAGACAAAGTACAGAT 4048
Qy 1422 ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGGCGAGGCTGGGCTCC 1481
Db 4049 ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGGCGAGGCTGGGCTCC 4108
Qy 1482 TTCTGCATGATTCTATCCAA 1501
Db 4109 TTCTGCATGATTCTATCCAA 4128

QY 222 TTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGGACCTATTGGAGCG 281
Db 2991 TTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGGACCTATTGGAGCG 3050
QY 282 ACTTTCCAGCAGTTTCAAGACAGAACGATGATACATAGGCGCTTCAAGAGGGAATTAAGAAA 341
Db 3051 ACTTTCCAGCAGTTTCAAGACAGAACGATGATACATAGGCGCTTCAAGAGGGAATTAAGAAA 3110
QY 342 CTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGC 401
Db 3111 CTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGC 3170
QY 402 CTTTGGAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTCTGAGAGAGAG 461
Db 3171 CTTTGGAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTCTGAGAGAGAG 3230
QY 462 CCCAGAAATGCTACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAAA 521
Db 3231 CCCAGAAATGCTACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAAA 3290
QY 522 AATTGAACCTGCACTCGCTGACTGGCAGAGAGAAAATAGATGAGACCTTGAAGACTCC 581
Db 3291 AATTGAACCTGCACTCGCTGACTGGCAGAGAGAAAATAGATGAGACCTTGAAGACTCC 3350
QY 582 AGGAATCTCAAGAGGCCAGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCA 641
Db 3351 AGGAATCTCAAGAGGCCAGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCA 3410
QY 642 AGGATCTGGAGCCGCTGGGAGTCTCTCTGACTCTCTCCAGATCTCTCCAGATCTCTGAGA 701
Db 3411 AGGATCTGGAGCCGCTGGGAGTCTCTCTGACTCTCTCCAGATCTCTCCAGATCTCTGAGA 3470
QY 702 AAGTCAAGGCACTTCGAGGAGAAATTTGGGCTCTGAAAGAGAACGCTGAGCCAGCTCAATG 761
Db 3471 AAGTCAAGGCACTTCGAGGAGAAATTTGGGCTCTGAAAGAGAACGCTGAGCCAGCTCAATG 3530
QY 762 ACCTTGCTGCGCACTTACACTTTGGGCAATGAGCTCTCAGCGTATTAACCTCAGACTC 821
Db 3531 ACCTTGCTGCGCACTTACACTTTGGGCAATGAGCTCTCAGCGTATTAACCTCAGACTC 3590
QY 822 TGGAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCGTGGAGACCGAGTCAAGC 881
Db 3591 TGGAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCGTGGAGACCGAGTCAAGC 3650
QY 882 AGCTGCATGAAGCCACAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTGTG 941
Db 3651 AGCTGCATGAAGCCACAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTGTG 3710
QY 942 TCCAGGTCCTGGAGAGACCACTCTGCCAACAAGTCCCTACTATATCAACCAAG 1001
Db 3711 TCCAGGTCCTGGAGAGACCACTCTGCCAACAAGTCCCTACTATATCAACCAAG 3770
QY 1002 AGACTCAAAACAATTTGCTGGGACCATCCAAAATGACAGAGCTTACCACTTTTAGCTG 1061
Db 3771 AGACTCAAAACAATTTGCTGGGACCATCCAAAATGACAGAGCTTACCACTTTTAGCTG 3830
QY 1062 ACCTGAAATAGTCCAGATTTCTAGCTTATAGGACTGCCATGAAACTCCGAGACTGAGA 1121
Db 3831 ACCTGAAATAGTCCAGATTTCTAGCTTATAGGACTGCCATGAAACTCCGAGACTGAGA 3890
QY 1122 AGGCCCTTTGCTGGATCTTTGAGCTGTGAGCTGTGATGATGCTTTGAGACAGACA 1181
Db 3891 AGGCCCTTTGCTGGATCTTTGAGCTGTGAGCTGTGATGATGCTTTGAGACAGACA 3950
QY 1182 ACCTCAAGCAAAATGACAGCCCATGGATATCTGAGATTAATTAATTTGTTGACCACTA 1241
Db 3951 ACCTCAAGCAAAATGACAGCCCATGGATATCTGAGATTAATTAATTTGTTGACCACTA 4010
QY 1242 TTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCTGGATA 1301
Db 4011 TTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCTGGATA 4070
QY 1302 TGTGCTGAAGCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGCTGTCC 1361

Db 4071 TGTGCTGAAGCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGCTGTCC 4130
QY 1362 TGTCTTTAAACCTGGCATCATTTCCCTGCTGTAAGACACATTTGGAAGACAGTACAGAT 1421
Db 4131 TGTCTTTAAACCTGGCATCATTTCCCTGCTGTAAGACACATTTGGAAGACAGTACAGAT 4190
QY 1422 ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTACCCAGCGCAGGCTGGGCTTCC 1481
Db 4191 ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTACCCAGCGCAGGCTGGGCTTCC 4250
QY 1482 TTCTGCATGATTTATCCAA 1501
Db 4251 TTCTGCATGATTTATCCAA 4270

RESULT 13
US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 5060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-36

Query Match 88.5%; Score 1328.8; DB 10; Length 5060;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 162 GGGTTCTTTTACAAGACAGTTCTGACCAAGTGAAGCGTGTGCACTTTCTTGCAGGAAC 221
Db 3001 GGTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGTGTGCACTTTCTTGCAGGAAC 3060
QY 222 TTCTGCTGGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGCACCCTATTGGAGCG 281
Db 3061 TTCTGCTGGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGCACCCTATTGGAGCG 3120
QY 282 ACTTTCCAGCAGTTTCAAGACAGAACGATGTACATAGGCGCTTCAAGAGGGAATTAAGAAA 341
Db 3121 ACTTTCCAGCAGTTTCAAGACAGAACGATGTACATAGGCGCTTCAAGAGGGAATTAAGAAA 3180
QY 342 CTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 401
Db 3181 CTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 3240
QY 402 CTTTGAAGGACTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTCTGAGGAGAG 461
Db 3241 CTTTGAAGGACTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTCTGAGGAGAG 3300
QY 462 CCCAGAAATGCTACTCGGCTTCTACGAAGCAGGCTGAGAGGTCAATATCTGAGTGGGAAA 521
Db 3301 CCCAGAAATGCTACTCGGCTTCTACGAAGCAGGCTGAGAGGTCAATATCTGAGTGGGAAA 3360
QY 522 AATTGAACCTGCACTCGCTGAGTGGCAGAGAAAATAGATGAGACCTTGAAGACTCC 581
Db 3361 AATTGAACCTGCACTCGCTGAGTGGCAGAGAAAATAGATGAGACCTTGAAGACTCC 3420
QY 582 AGGAATCTCAAGAGGCCAGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCA 641
Db 3421 AGGAATCTCAAGAGGCCAGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCA 3480

QY 642 AGGATCTGGCAGCCGCTGGCGGATCTCTCTCAAGATCACCTCGAGA 701
Db 3481 AGGATCTGGCAGCCGCTGGCGGATCTCTCTCAAGATCACCTCGAGA 3540
QY 702 AAGTCAAGGCACTTCGAGGAGAAATGGCGCTCTGAAAGAGAAAGTGAAGCAAGTCAATG 761
Db 3541 AAGTCAAGGCACTTCGAGGAGAAATGGCGCTCTGAAAGAGAAAGTGAAGCAAGTCAATG 3600
QY 762 ACCTTGCTCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATTAACCTCAGCACTC 821
Db 3601 ACCTTGCTCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATTAACCTCAGCACTC 3660
QY 822 TGAAGACCTGAAACACCAAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCAAGC 881
Db 3661 TGAAGACCTGAAACACCAAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCAAGC 3720
QY 882 AGCTGCATGAGCCACAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTG 941
Db 3721 AGCTGCATGAGCCACAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTG 3780
QY 942 TCCAGGCTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCAAG 1001
Db 3781 TCCAGGCTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCAAG 3840
QY 1002 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTTACCACTCTTTAGCTG 1061
Db 3841 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTTACCACTCTTTAGCTG 3900
QY 1062 ACCTGAATTAATGTCAGATTTCTCAGCTTATAGGCTGCCATGAACTCCGAGAGCTCGA 1121
Db 3901 ACCTGAATTAATGTCAGATTTCTCAGCTTATAGGCTGCCATGAACTCCGAGAGCTCGA 3960
QY 1122 AGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCAAGCA 1181
Db 3961 AGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGCTGCATGTGATGCCCTTGGACCAAGCA 4020
QY 1182 ACCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATTAATTTGTTGACCACTA 1241
Db 4021 ACCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATTAATTTGTTGACCACTA 4080
QY 1242 TTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGCTCCCTCTCTCGCTGGATA 1301
Db 4081 TTTATGACCGCTGGAGCAAGAGCACAAATTTGGTCAACGCTCCCTCTCTCGCTGGATA 4140
QY 1302 TGTCTGTAAGTCTGCTGATTTTATGATAGGAGCAAGAGGAGATCCGTGCTC 1361
Db 4141 TGTCTGTAAGTCTGCTGATTTTATGATAGGAGCAAGAGGAGATCCGTGCTC 4200
QY 1362 TGTCTTTAAACCTGGCATCATTTCCCTGTGTAAGCACATTTGGAGACAAAGTACAGAT 1421
Db 4201 TGTCTTTAAACCTGGCATCATTTCCCTGTGTAAGCACATTTGGAGACAAAGTACAGAT 4260
QY 1422 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAAGCGAGCTGGGCTCC 1481
Db 4261 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAAGCGAGCTGGGCTCC 4320
QY 1482 TTTGTCATGATTTCTATCCAA 1501
Db 4321 TTTGTCATGATTTCTATCCAA 4340

RESULT 14

US-09-845-416-27
; Sequence 27, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777

; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 5149
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-27

Query Match 88.5%; Score 1328.8; DB 10; Length 5149;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 162 GGGTTCTTTTCAACACACAGTTCTGACCACTGGAAGCGTCTCACCTTTCTGCGAGGAAC 221
Db 3090 GGTCCATTGGAAGCCAGTTCTGACCACTGGAAGCGTCTGACCTTTCTGCGAGGAAC 3149
QY 222 TTCTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGCACCTATTGGAGGCG 281
Db 3150 TTCTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGCACCTATTGGAGGCG 3209
QY 282 ACTTTCCAGCAGTTCCAGAGCAGACGATGATACATAGGCGCTTCAAGAGGGAATTTGAAA 341
Db 3210 ACTTTCCAGCAGTTCCAGAGCAGACGATGATACATAGGCGCTTCAAGAGGGAATTTGAAA 3269
QY 342 CTAAAGAAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 401
Db 3270 CTAAAGAAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 3329
QY 402 CTTTGGAAAGGACTAGAGAAACTCTACAGAGAGCCGAGAGAGCTGCTCCTGAGGAGAGAG 461
Db 3330 CTTTGGAAAGGACTAGAGAAACTCTACAGAGAGCCGAGAGAGCTGCTCCTGAGGAGAGAG 3389
QY 462 CCCAGAAATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAAA 521
Db 3390 CCCAGAAATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAAA 3449
QY 522 AATTGAAACCTGCACTCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTTGAAAGACTCC 581
Db 3450 AATTGAAACCTGCACTCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTTGAAAGACTCC 3509
QY 582 AGGAATTTCAAGAGCCACGATGAGCTGGAGCTCAAGCTGCGCCAGAGCTGAGGTGATCA 641
Db 3510 AGGAATTTCAAGAGCCACGATGAGCTGGAGCTCAAGCTGCGCCAGAGCTGAGGTGATCA 3569
QY 642 AGGATCTCGCAGCCCGTGGCGATCTCTCTCAATTTGACTCTCTCCAAGATCACCTCGAGA 701
Db 3570 AGGATCTCGCAGCCCGTGGCGATCTCTCTCAATTTGACTCTCTCCAAGATCACCTCGAGA 3629
QY 702 AAGTCAAGGCACTTCGAGGAGAAATTTGGCGCTCTGAAAGAGAAAGTGAAGCAAGTCAATG 761
Db 3630 AAGTCAAGGCACTTCGAGGAGAAATTTGGCGCTCTGAAAGAGAAAGTGAAGCAAGTCAATG 3689
QY 762 ACCTTGCTCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAAGTCAAGCACTC 821
Db 3690 ACCTTGCTCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAAGTCAAGCACTC 3749
QY 822 TGAAGACCTGAAACACCAAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCAAGC 881
Db 3750 TGAAGACCTGAAACACCAAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCAAGC 3809
QY 882 AGCTGCATGAGCCACAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTG 941
Db 3810 AGCTGCATGAGCCACAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTG 3869
QY 942 TCCAGGCTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCAAG 1001
Db 3870 TCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCAAG 3929
QY 1002 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTAGCTG 1061
Db 3930 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTAGCTG 3989

Qy	1422	ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCGCAGGCTGGGCTCC	1481
Db	3135	ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCGCAGGCTGGGCTCC	3194
Qy	1482	TTCTGCATGATTCATCAA	1501
Db	3195	TTCTGCATGATTCATCAA	3214

Search completed: September 19, 2004, 15:14:18
Job time : 503.133 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2004, 07:39:06 ; Search time 543.946 Seconds
(without alignments)
15627.746 Million cell updates/sec

Title: US-09-845-416-14_COPY_1000_3000
Perfect score: 2001
Sequence: 1 gscagtcattgatgagag.....cattgagccaagtgtccgga 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002s.*
- 7: Geneseqn2003as.*
- 8: Geneseqn2003bs.*
- 9: Geneseqn2003cs.*
- 10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2201	100.0	3446	6	AAD37242	Aad37242 Human dys
2	1990	99.5	4414	6	AAD37260	Aad37260 Adeno-ass
3	1820.8	91.0	5417	6	ABK81997	Abk81997 DNA encod
4	1593	79.6	3510	6	AAD37240	Aad37240 Human dys
5	1593	79.6	4476	6	AAD37259	Aad37259 Adeno-ass
6	1332	66.6	4402	3	AZ48568	Az48568 A rod sho
7	1269.2	63.4	4402	3	AZ48567	Az48567 A rod sho
8	1263	63.1	3858	6	AAD37237	Aad37237 Human dys
9	1263	63.1	4825	6	AAD37257	Aad37257 Adeno-ass
10	1263	63.1	4848	6	AAD37263	Aad37263 Adeno-ass
11	1263	63.1	5060	6	AAD37264	Aad37264 Adeno-ass
12	1254	62.7	5339	6	ABK81998	Abk81998 DNA encod
13	1251.2	62.5	5331	6	AAD37238	Aad37238 Human dys
14	1251.2	62.5	4498	6	AAD37258	Aad37258 Adeno-ass
15	1245	62.2	4182	6	AAD37230	Aad37230 Human dys
16	1245	62.2	5149	6	AAD37255	Aad37255 Adeno-ass
17	1132.2	56.6	5462	6	ABK81999	Abk81999 DNA encod
18	1112	55.6	3999	6	AAD37234	Aad37234 Human dys
19	1112	55.6	4966	6	AAD37256	Aad37256 Adeno-ass
20	1112	55.6	4990	6	AAD37262	Aad37262 Adeno-ass
21	1002.6	50.1	1821	6	AAD37241	Aad37241 Human dys
22	1002.6	50.1	2169	6	AAD37232	Aad37232 Human dys
23	1002.6	50.1	5952	5	AAD06794	Aad06794 Human dys

24	1002.6	50.1	8689	6	ABK82000	Abk82000 DNA encod
25	1002.6	50.1	11058	6	AAD37229	Aad37229 Human dys
26	1002.6	50.1	11241	6	ABK82005	Abk82005 CDNA enco
27	1002.6	50.1	11443	6	ABK82002	Abk82002 DNA encod
28	1002.6	50.1	12923	1	AA90338	Aan90338 Sequence
29	1002.6	50.1	13957	6	ABK81959	Abk81959 CDNA enco
30	1002.6	50.1	13957	6	ABT10904	Abt10904 Human bre
31	1002.6	50.1	13957	6	ABN95786	Abn95786 Gene #228
32	1002.6	50.1	13957	6	ABS69900	Abs69900 Human dys
33	1001	50.0	1434	6	AAD37243	Aad37243 Human dys
34	991.6	49.6	13977	6	ABS70403	AbS70403 Human bon
35	985.8	49.3	13991	6	AAD37231	Aad37231 Human dys
36	869.8	43.5	13815	6	ABK81960	Abk81960 CDNA enco
37	869.8	43.5	13815	6	AB193799	Ab199799 Mouse isc
38	869.8	43.5	13815	6	AAT27558	Aat27558 Shuttle v
39	868.2	43.4	13815	2	AAV18885	AAV18885 Mus muscu
40	767	38.3	3275	1	AA97129	Aan97129 Partial s
41	735.4	36.8	4075	3	AZ48569	Az48569 A rod sho
42	731.2	36.5	3747	3	AZ48566	Az48566 A rod sho
43	677.6	33.9	3163	3	AZ48571	Az48571 A rod sho
44	665	33.2	1667	6	AAD37235	Aad37235 Human dys
45	567.4	28.4	2005	7	ABT33376	Abt33376 NOVX DNA

ALIGNMENTS

RESULT 1
AAD37242
ID AAD37242 standard; DNA; 3446 BP.
XX
AC AAD37242;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3447.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WC200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
XX (XIAO/) XIAO X.
XX
Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 53-54; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3447 containing nucleotides 1-1992 (N-terminus,
CC hinge H1 and rods R1, R2 and R3), 8749-10227 (rod R24, hinge H4 and CR
CC domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ Sequence 3446 BP; 1074 A; 766 C; 787 G; 819 T; 0 U; 0 Other;

Query Match 100.0%; Score 2001; DB 6; Length 3446;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 60
Db 1000 GCGAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 1059

QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTCTTAAT 120
Db 1060 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTCTTAAT 1119

QY 121 GATGTGAAGTGTGCAAGAGACCACTTTTCATCTCATGAGGGTACATGATGGATTTGACA 180
Db 1120 GATGTGAAGTGTGCAAGAGACCACTTTTCATCTCATGAGGGTACATGATGGATTTGACA 1179

QY 181 GCCCATCAGGCCCGGGTGTGTAATATTCATAATTTGGAAAGTAAAGCTGATGGAACAGGA 240
Db 1180 GCCCATCAGGCCCGGGTGTGTAATATTCATAATTTGGAAAGTAAAGCTGATGGAACAGGA 1239

QY 241 AAATTATCAGAAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGA 300
Db 1240 AAATTATCAGAAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGA 1299

QY 301 TGGGAATGCTTCAGGTAGCTAGCATGGAAAAAACAAGCAATTTACATAGAGTTTAAATG 360
Db 1300 TGGGAATGCTTCAGGTAGCTAGCATGGAAAAAACAAGCAATTTACATAGAGTTTAAATG 1359

QY 361 GATCTCCAGAACTGAAACTGAAGAGTTGAATGACTGGCTTAACAAAAACAAGAAAGAA 420
Db 1360 GATCTCCAGAACTGAAACTGAAGAGTTGAATGACTGGCTTAACAAAAACAAGAAAGAA 1419

QY 421 CAAGGAAATGGAGAAAGACCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTAC 480
Db 1420 CAAGGAAATGGAGAAAGACCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTAC 1479

QY 481 AACACATAGGTGCTTCAAGAGATCTAGAAACAAGTCAAGGTGCAATTTCTCTCA 540
Db 1480 AACACATAGGTGCTTCAAGAGATCTAGAAACAAGTCAAGGTGCAATTTCTCTCA 1539

QY 541 CTCACATGGTGGTGGTGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAG 600
Db 1540 CTCACATGGTGGTGGTGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAG 1599

QY 601 AACAACTTAAGGTATTGGGAGATCGATGGCAAAACATCTGTAGATGGAACAAGACCGCT 660
Db 1600 AACAACTTAAGGTATTGGGAGATCGATGGCAAAACATCTGTAGATGGAACAAGACCGCT 1659

QY 661 GGGTCTTTTACAAGACATCTTCTCAAAATGGCAAGCTCTTACTGAAGAACAGTGCCTTT 720
Db 1660 GGGTCTTTTACAAGACATCTTCTCAAAATGGCAAGCTCTTACTGAAGAACAGTGCCTTT 1719

QY 721 TTAGTGCATGGCTTTCAAAAAAGAAGATGCAAGTGAACAAGATTCACAACTGGCTTTA 780
Db 1720 TTAGTGCATGGCTTTCAAAAAAGAAGATGCAAGTGAACAAGATTCACAACTGGCTTTA 1779

QY 781 AAGATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGCGCTTTAAAGCGGATCTAG 840
Db 1780 AAGATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGCGCTTTAAAGCGGATCTAG 1839

QY 841 AAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTCAAAACAAGATCTCTTTCAACAC 900
Db 1840 AAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTCAAAACAAGATCTCTTTCAACAC 1899

QY 901 TGAAGATTAAGTCAAGTGAACCCAGAACCGAAGCATGGCTGGATTAATCTTGGCCCGTGT 960
Db 901 TGAAGATTAAGTCAAGTGAACCCAGAACCGAAGCATGGCTGGATTAATCTTGGCCCGTGT 960

Db 1900 TGAAGAAATAGTCAAGTGAAGCCAGGAAAGACGGAAGCATGGCTGGATAACTTTCCCGCGTGT 1959
QY 961 GGGATAAATTTAGTCCAAAAAAGTGTGAAAAGAGTACAGACAGACCCCTTGAAGAGCTCCAGG 1020
Db 1960 GGGATAAATTTAGTCCAAAAAAGTGTGAAAAGAGTACAGACAGACCCCTTGAAGAGCTCCAGG 2019

QY 1021 AACTTCAAGAGCCACGGATGAGCTGGACCTCAAGCTGGCCAGCTGAGTGATCAAGG 1080
Db 2020 AACTTCAAGAGCCACGGATGAGCTGGACCTCAAGCTGGCCAGCTGAGTGATCAAGG 2079

QY 1081 GATCTCTGGCAGCCCGTGGCGGATCTCTCATTTGATCTCTCTCAAGATCACTCTCGAGAAG 1140
Db 2080 GATCTCTGGCAGCCCGTGGCGGATCTCTCATTTGATCTCTCTCCAGATCACTCTCGAGAAG 2139

QY 1141 TCAAGGCATCTTGAGGAGAAATTTGGCCCTCTGAAAGAGAAAGTGAAGCCACGTCAATGACC 1200
Db 2140 TCAAGGCATCTTGAGGAGAAATTTGGCCCTCTGAAAGAGAAAGTGAAGCCACGTCAATGACC 2199

QY 1201 TTGCTCGCCAGCTTACCACTTTTGGCATTTTGGCATTTTGGCATTTTGGCATTTTGGCATTTG 1260
Db 2200 TTGCTCGCCAGCTTACCACTTTTGGCATTTTGGCATTTTGGCATTTTGGCATTTTGGCATTTG 2259

QY 1261 AAGACCTGAAACACAGATGGAAGCTTTTGCAGGTGGCGCTCGAGGACCGAGTCAGGACG 1320
Db 2260 AAGACCTGAAACACAGATGGAAGCTTTTGCAGGTGGCGCTCGAGGACCGAGTCAGGACG 2319

QY 1321 TGCATGAGCCCAACAGGACCTTTTGGTCCAGCATCTCAGCATCTTCTTTCACAGTCTCTCC 1380
Db 2320 TGCATGAGCCCAACAGGACCTTTTGGTCCAGCATCTCAGCATCTTCTTTCACAGTCTCTCC 2379

QY 1381 AGGCTCCCTGGGAGAGACCATCTCGCCAAACAAAGTGCCCTACTATATCAACACACAGA 1440
Db 2380 AGGCTCCCTGGGAGAGACCATCTCGCCAAACAAAGTGCCCTACTATATCAACACACAGA 2439

QY 1441 CTCAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACC 1500
Db 2440 CTCAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACC 2499

QY 1501 TGAATAATGTGAGATTTCTCAGCTTATAGGACTGCGCATGAAACTCCGAAGACTGCAGAAAG 1560
Db 2500 TGAATAATGTGAGATTTCTCAGCTTATAGGACTGCGCATGAAACTCCGAAGACTGCAGAAAG 2559

QY 1561 CCTTTGGTTGGATCTCTTGAGCCTGTGAGCTGATGATGCTTGGACAGCAACCAACC 1620
Db 2560 CCTTTGGTTGGATCTCTTGAGCCTGTGAGCTGATGATGCTTGGACAGCAACCAACC 2619

QY 1621 TCAAGCAAAATGACACCGCCATGATATCTGAGATTTATTAATTTGTTGACCACTATTT 1680
Db 2620 TCAAGCAAAATGACACCGCCATGATATCTGAGATTTATTAATTTGTTGACCACTATTT 2679

QY 1681 ATGACCCCTGGAGCAAGAGCAACAAATTTGGTCAACGCTCCCTCTCTGCTGGATATGT 1740
Db 2680 ATGACCCCTGGAGCAAGAGCAACAAATTTGGTCAACGCTCCCTCTCTGCTGGATATGT 2739

QY 1741 GTCTGAATCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGCTGCTCTGT 1800
Db 2740 GTCTGAATCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGCTGCTCTGT 2799

QY 1801 CTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACTATTTGGAAGCAAGTACAGATACC 1860
Db 2800 CTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACTATTTGGAAGCAAGTACAGATACC 2859

QY 1861 TTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACAGCCAGGCTGGGCTCTCTTC 1920
Db 2860 TTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACAGCCAGGCTGGGCTCTCTTC 2919

QY 1921 TGCATGATTTCTATCCAAATTCGAAGCAGTGGGTGAAGTTGATCTCTTTGGGGGAGTA 1980
Db 2920 TGCATGATTTCTATCCAAATTCGAAGCAGTGGGTGAAGTTGATCTCTTTGGGGGAGTA 2979

QY 1981 ACATTGAGCAAGTGTCCGGA 2001
Db 2980 ACATTGAGCAAGTGTCCGGA 3000

RESULT 2
AAD37260
ID AAD37260 standard; DNA; 4414 BP.
XX
AC AAD37260;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-3447.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
XX WO200183695-A2.
XX
XX 08-NOV-2001.
XX
XX 27-APR-2001; 2001WO-US013677.
XX
XX 28-APR-2000; 2000US-0200777P.
XX
XX (XIAO/) XIAO X.
XX
XX Xiao X;
XX
XX WFI; 2002-049342/06.
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
XX Example 1; Page 65-66; 71pp; English.
XX
XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ Sequence 4414 BP; 1255 A; 1075 C; 1086 G; 998 T; 0 U; 0 Other;

Query Match 99.5%; Score 1990; DB 6; Length 4414;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTTAGAAGAA 60
Db 1757 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTTAGAAGAA 1816

QY 61 GTATTATCGTGGCTTCTTTCTGCTGGGACACATTCGAAGCACAAGAGAGATTTCTAAT 120
Db 1817 GTATTATCGTGGCTTCTTTCTGCTGGGACACATTCGAAGCACAAGAGAGATTTCTAAT 1876

QY 121 GATGTGAAGTGGTGAAGACCAGTTTCATACCTCATGAGGGGTACATGATGGATTTGACA 180
Db 1877 GATGTGAAGTGGTGAAGACCAGTTTCATACCTCATGAGGGGTACATGATGGATTTGACA 1936

QY 181 GCCCATCAGGCCGGGTGGTAATATCTACAAATGGGAAGTAAAGTGGATGGAAACAGGA 240

Db 1937 GCCCATCAGGCCGGGTGGTAATATCTCAATTTGGGAATAGCTGATTTGGAAACAGGA 1996
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTCAAGAGCAGATGAATCTCTTAAATTCAGA 300
Db 1997 AAATTATCAGAAGATGAAGAACTGAAGTCAAGAGCAGATGAATCTCTTAAATTCAGA 2056
QY 301 TGGGAATGCCTCAGGCTAGCTAGCATGGAAAAAACAAGCAATTTACATAGAGTTTTAATG 360
Db 2057 TGGGAATGCCTCAGGCTAGCTAGCATGGAAAAAACAAGCAATTTACATAGAGTTTTAATG 2116
QY 361 GATCTCCAGATC-GAAACTGAAAGAGTTGAATCACTGGCTTAAACAAAAACAGAAAGAA 419
Db 2117 GATCTCCAGATCAGAACTGAAAGAGTTGAATCACTGGCTTAAACAAAAACAGAAAGAA 2176
QY 420 ACAAGGAAATGGAGGAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACGCAAGTA 479
Db 2177 ACAAGGAAATGGAGGAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACGCAAGTA 2236
QY 480 CAACAACATTAAGTGTCTTCAAGAGATCTAGAAACAAGCAAGTCAGGGTCAATCTCTC 539
Db 2237 CAACAACATTAAGTGTCTTCAAGAGATCTAGAAACAAGTCAGGGTCAATCTCTC 2296
QY 540 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGTCTTTGGAA 599
Db 2297 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGTCTTTGGAA 2356
QY 600 GAACAACCTTAAGTATTGGGAGATCGATGGCAACATCTCTAGATGGACAGAGACCGC 659
Db 2357 GAACAACCTTAAGTATTGGGAGATCGATGGCAACATCTCTAGATGGACAGAGACCGC 2416
QY 660 TGGGTTCTTTTACAAAGACATCCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTT 719
Db 2417 TGGGTTCTTTTACAAAGACATCCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTT 2476
QY 720 TTTAGTCATGGCTTTTCAAAAAAGAGATCGATGAAACAAGATTCACACAACCTGGCTTT 779
Db 2477 TTTAGTCATGGCTTTTCAAAAAAGAGATCGATGAAACAAGATTCACACAACCTGGCTTT 2536
QY 780 AAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTA 839
Db 2537 AAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTA 2596
QY 840 GAAAAAGAAAAAGCAATCCATGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACA 899
Db 2597 GAAAAAGAAAAAGCAATCCATGGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACA 2656
QY 900 CTGAAGATAATAGTCAAGTCAACCCAGAGACGGAAGCTGGTGGATACTTTGCCCGGTGT 959
Db 2657 CTGAAGATAATAGTCAAGTCAACCCAGAGACGGAAGCTGGTGGATACTTTGCCCGGTGT 2716
QY 960 TGGGATAATTTAGTCCAAAAAATTTGAAAAGAGTACAGCACAGACCTTGAAGAGACTCCAG 1019
Db 2717 TGGGATAATTTAGTCCAAAAAATTTGAAAAGAGTACAGCACAGACCTTGAAGAGACTCCAG 2776
QY 1020 GAACTTCAAGAGGCCACCGATGAGTGAACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAG 1079
Db 2777 GAACTTCAAGAGGCCACCGATGAGTGAACCTCAAGCTGCGCCCAAGCTGAGGTGATCAAG 2836
QY 1080 GGATCCTGCGACCCGCTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAA 1139
Db 2837 GGATCCTGCGACCCGCTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAGAAA 2896
QY 1140 GTCAAGGCACTTCGAGGAGAAAATTCGGCTCTGAAAAGAGAACGTGAGCCACGTCAATGAC 1199
Db 2897 GTCAAGGCACTTCGAGGAGAAAATTCGGCTCTGAAAAGAGAACGTGAGCCACGTCAATGAC 2956
QY 1200 CTTGCTGCGACCTTACCACTTTGGGCATTGAGTCTCAGCTCTCACCCTTAACTCAGACTCTG 1259
Db 2957 CTTGCTGCGACCTTACCACTTTGGGCATTGAGTCTCAGCTCTCACCCTTAACTCAGACTCTG 3016
QY 1260 GAAGACCTGAAACACACAGATGGAAGTTCTGAGGTGGCCGCTCGAGGACCGGAGTCAGGCAG 1319

Db 3017 GAAGACCTGTAACACCATGATGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGCAG 3076
Qy 1320 CTGCATGAAGCCACACAGGACTTTGGTCCAGCATCTCTCAGCACTTTCTTCCAGTCTGTC 1379
Db 3077 CTGCATGAAGCCACACAGGACTTTGGTCCAGCATCTCTCAGCACTTTCTTCCAGTCTGTC 3136
Qy 1380 CAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACGAG 1439
Db 3137 CAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACGAG 3196
Qy 1440 ACTCAAAACACTTCTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGAC 1499
Db 3197 ACTCAAAACACTTCTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGAC 3256
Qy 1500 CTGAATATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCGAAG 1559
Db 3257 CTGAATATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCGAAG 3316
Qy 1560 GCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGCTGATGCTTGGACGACAC 1619
Db 3317 GCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGCTGATGCTTGGACGACAC 3376
Qy 1620 CTCAAGCAAAATGACCAAGCCCATGGATATCTCTCAGATTATTAATTGTTGACCACTAT 1679
Db 3377 CTCAAGCAAAATGACCAAGCCCATGGATATCTCTCAGATTATTAATTGTTGACCACTAT 3436
Qy 1680 TATGACCCCTGGAGCAAGAGACACAAATTTGGTCAACGCTCCTCTCTGCTGGATG 1739
Db 3437 TATGACCCCTGGAGCAAGAGACACAAATTTGGTCAACGCTCCTCTCTGCTGGATG 3496
Qy 1740 TGTCTGAACCTGGCTGCTGAATGTTTATGATCGGACGAACTGGAGGATCGGTGCTG 1799
Db 3497 TGTCTGAACCTGGCTGCTGAATGTTTATGATCGGACGAACTGGAGGATCGGTGCTG 3556
Qy 1800 TCTTTTAAACCTGGATCTTCCCTGTGTAAAGCACTTTGGAACAAAGTACAGATAC 1859
Db 3557 TCTTTTAAACCTGGATCTTCCCTGTGTAAAGCACTTTGGAACAAAGTACAGATAC 3616
Qy 1860 CTTTTCAGCAAGTGGCAAGTCAACAGGATTTTGTGACAGCGAGGCTGGGCTCCTT 1919
Db 3617 CTTTTCAGCAAGTGGCAAGTCAACAGGATTTTGTGACAGCGAGGCTGGGCTCCTT 3676
Qy 1920 CTGCATGATTCTATCCAAATTCGAAGACAGTGGGTGAAGTTGATCTCTTTGGGGCAGT 1979
Db 3677 CTGCATGATTCTATCCAAATTCGAAGACAGTGGGTGAAGTTGATCTCTTTGGGGCAGT 3736
Qy 1980 AACATTGAGCCAGTGTCCGGA 2001
Db 3737 AACATTGAGCCAGTGTCCGGA 3758

RESULT 3

ID ABK81997 standard; DNA; 5417 BP.

XX AC ABK81997;

XX DT 13-AUG-2002 (first entry)

XX DE DNA encoding mini-dystrophin protein deltaR4-R23.

XX KN Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;

XX KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN W0200229056-A2.

XX PX 11-APR-2002.

XX PF 04-OCT-2001; 2001WO-US031126.

XX XX

PR 06-OCT-2000; 2000US-0238848P.
XX (UNMI) UNIV MICHIGAN.
XX Chamberlain JS, Harper SQ;
XX WPI; 2002-435334/46.
DR A composition for preparing therapeutic drugs, has a mini-dystrophin
XX peptide comprising a specific number of spectrin-like repeat domains, or
PT a nucleic acid sequence encoding the mini-dystrophin peptide.
XX Disclosure; Fig 12; 145pp; English.
XX The invention describes a composition comprising a mini-dystrophin
XX peptide comprising a spectrin-like repeat domain, where the domain
XX comprises n spectrin-like repeats, and contains no more than n spectrin-
XX like repeats, where n is an even number between 4-24, or a nucleic acid
XX encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
XX polynucleotide encoding it is useful as a medicament, for preparing a
XX drug for therapeutic application and in the preparation of a composition
XX for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
XX (DMD). This sequence represents a mini-dystrophin sequence of the
XX invention
XX Sequence 5417 BP; 1700 A; 1192 C; 1182 G; 1343 T; 0 U; 0 Other;

Query Match 91.0%; Score 1820.8; DB 6; Length 5417;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 2; Indels 157; Gaps 2;

Qy 1 GCGAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 60
Db 1199 GCGAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 1258
Qy 61 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTCGAAGCAAGAGAGAGATTTCTAAT 120
Db 1259 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTCGAAGCAAGAGAGATTTCTAAT 1318
Qy 121 GATGTGAAGTGTGAAAGACCCAGTTTCATCTACTCATGAGGGTACATGATGGATTGACA 180
Db 1319 GATGTGAAGTGTGAAAGACCCAGTTTCATCTACTCATGAGGGTACATGATGGATTGACA 1378
Qy 181 GCCCATCAGGCCCGGGTGGTAAATTTCTACAATTTGGGAAAGTAAAGCTGATTGAAACAGGA 240
Db 1379 GCCCATCAGGCCCGGGTGGTAAATTTCTACAATTTGGGAAAGTAAAGCTGATTGAAACAGGA 1438
Qy 241 AAATTATCAGAGATGAAGAACTGAAAGTACAGAGCAGATGAATCTCTTAATTTCAAGA 300
Db 1439 AAATTATCAGAGATGAAGAACTGAAAGTACAGAGCAGATGAATCTCTTAATTTCAAGA 1498
Qy 301 TGGGAATGCTCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGATTTTAATG 360
Db 1499 TGGGAATGCTCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGATTTTAATG 1558
Qy 361 GATCTCCAGAAATC-GAAACTGAAAGAGTTGAATGACTGGCTAAACAAAGCAAGAAAGA 419
Db 1559 GATCTCCAGAAATCAGAAATCTGAAAGAGTTGAATGACTGGCTAAACAAAGCAAGAAAGA 1618
Qy 420 ACAGGGAATGAGGAGAGAGCTCTTGGACCTGATCTTGAAGCACTTAAACCCCAAGTA 479
Db 1619 ACAGGGAATGAGGAGAGAGCTCTTGGACCTGATCTTGAAGCACTTAAACCCCAAGTA 1678
Qy 480 CAACAACATAGGTGCTTCAAGAGATCTAGAACAAAGAACCAAGTCAGGGTCAATTTCTCT 539
Db 1679 CAACAACATAGGTGCTTCAAGAGATCTAGAACAAAGAACCAAGTCAGGGTCAATTTCTCT 1738
Qy 540 ACTCAGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 599
Db 1739 ACTCAGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1798
Qy 600 GAACAACCTTAAGGTATTGGGAGATCGATGGCAAAACATCTGTAGATGGACAGAGACCCG 659

Db 1799 GRACAACTTAAGGATTTGGAGATCGATGGGCAACATCTGTAGATGGACAGAGACCGC 1858
Qy 660 TGGGTTCTTTTAAAGACATCTCTCAAAATGCGAAGCTCTTACTGAAAGAACAGTGCCTTT 719
Db 1859 TGGGTTCTTTTAAAGACATCTCTCAAAATGCGAAGCTCTTACTGAAAGAACAGTGCCTTT 1918
Qy 720 TTTAGTGCATGGCTTTTCAAGAAAAGAGATGCAAGTGAACAAGATTCACACAACTGGCTTT 779
Db 1919 TTTAGTGCATGGCTTTTCAAGAAAAGAGATGCAAGTGAACAAGATTCACACAACTGGCTTT 1978
Qy 780 AAGAGTCAAAATGAAATGTTATCAAGTCTTTCAAAAATCTGGCCGTTTAAAAAGCGGATCTA 839
Db 1979 AAGAGTCAAAATGAAATGTTATCAAGTCTTTCAAAAATCTGGCCGTTTAAAAAGCGGATCTA 2038
Qy 840 GAAAGAAAAGCAATTCATGGGCAAACTGTATTCATCTCAAAAGATCTCTTTTCAACA 899
Db 2039 GAAAGAAAAGCAATTCATGGGCAAACTGTATTCATCTCAAAAGATCTCTTTTCAACA 2098
Qy 900 CTGAAGAATTAAGTCAGTGAACCGAAGACGGAAGCATGGTGTGATTAACCTTTGGCCCGGTGT 959
Db 2099 CTGAAGAATTAAGTCAGTGAACCGAAGACGGAAGCATGGTGTGATTAACCTTTGGCCCGGTGT 2158
Qy 960 TGGGATTAATTTAGTCCAAAACCTTTGAAAAGAGTACAGCACAGA----- 1002
Db 2159 TGGGATTAATTTAGTCCAAAACCTTTGAAAAGAGTACAGCACAGATTTTACAGGCGTGCACC 2218
Qy 1003 ----- 1002
Db 2219 ACCACTCAGCCATCACTAACACAGACAACTGTATTTGGAACACAGTAACTACGCTGACCCACA 2278
Qy 1003 ----- 1002
Db 2279 AGGGAACAGATCCTGTGTAAGCATGCTCAAGAGGAACCTTCCACCAACCATCCGCCAAAAG 2338
Qy 1003 -----CCCTTGAAGACTCCAGGAACCTTCAAGAGCCACCGATGAG 1043
Db 2339 AAGAGGAGATTAAGTGTGATCTTTGAAAGACTCCAGGAACCTTCAAGAGCCACCGATGAG 2398
Qy 1044 CTGACCTCAAGCTGCGCCAAGTGAAGTGTATCAAGGGATCTCTGGCAGCCCGCTGGCGCAT 1103
Db 2399 CTGACCTCAAGCTGCGCCAAGTGAAGTGTATCAAGGGATCTCTGGCAGCCCGCTGGCGCAT 2458
Qy 1104 CTCTCATTAAGTCTCTCCAGATCACTCGAAGAGTCAAGGACTTCGAGAGAAATTT 1163
Db 2459 CTCTCATTAAGTCTCTCCAGATCACTCGAAGAGTCAAGGACTTCGAGAGAAATTT 2518
Qy 1164 GCGCTCTGAAAGAGAACGTCGAGCCAGCTCAATGACCTTCTGCCAGCTTACCACTTTG 1223
Db 2519 GCGCTCTGAAAGAGAACGTCGAGCCAGCTCAATGACCTTCTGCCAGCTTACCACTTTG 2578
Qy 1224 GGCATTCAGCTCTCAACCGTATAACCTCAGCACTCTGGAAGACCTGAAACCCAGATGGAAG 1283
Db 2579 GGCATTCAGCTCTCAACCGTATAACCTCAGCACTCTGGAAGACCTGAAACCCAGATGGAAG 2638
Qy 1284 CTTCTGAGGTGGCGTCGAGGCCGAGTCAGCAGCTGATGAAGCCACAGGGACTTT 1343
Db 2639 CTTCTGAGGTGGCGTCGAGGCCGAGTCAGCAGCTGATGAAGCCACAGGGACTTT 2698
Qy 1344 GGTCCAGCATCTCAGCACTTTCTTTCCACGCTGTCTCCAGGGTCTCCGGGAGAGAGCCATC 1403
Db 2699 GGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTCTCCAGGGTCTCCGGGAGAGAGCCATC 2758
Qy 1404 TCGCCAAACAAAGTGGCCCTACTATATCAACACAGAGACTCAAAACACTTGTCTGGGACCAT 1463
Db 2759 TCGCCAAACAAAGTGGCCCTACTATATCAACACAGAGACTCAAAACACTTGTCTGGGACCAT 2818
Qy 1464 CCCAAAATGACAGAGCTCTACCACTCTTTAGCTTGACCTGAAATATGTCAGATTTCTCAGCT 1523
Db 2819 CCCAAAATGACAGAGCTCTACCACTCTTTAGCTTGACCTGAAATATGTCAGATTTCTCAGCT 2878
Qy 1524 TATAGGACTGCAATGAAACTCCGAAAGACTGCAAGAGCCCTTTGCTTGGATCTCTTTGAGC 1583
Db 2879 TATAGGACTGCAATGAAACTCCGAAAGACTGCAAGAGCCCTTTGCTTGGATCTCTTTGAGC 2938

RESULT 4

AAD37240

ID AAD37240 standard; DNA; 3510 BP.

XX AAD37240;

XX AAD37240;

XX DT 21-AUG-2002 (first entry)

XX DE Human dystrophin minigene delta3510.

XX KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

XX KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX KW Becker muscular dystrophy; ds.

XX OS Homo sapiens.

XX PN WO200183695-A2.

XX PD 08-NOV-2001.

XX PP 27-APR-2001; 2001WO-US013677.

XX PR 28-APR-2000; 2000US-0200777P.

XX PA (XIAO/) XIAO X.

XX PI Xiao X;

XX DR WPI; 2002-049342/06.

XX PT New dystrophin minigene for treating Duchenne or Becker muscular

XX PT dystrophy comprises an N-terminal domain or modified N-terminal domain,

XX PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin

XX PT gene.

XX PS Example 1; Page 51-52; 71pp; English.

XX CC The present invention relates to an isolated nucleotide sequence encoding

XX CC a dystrophin minigene. The minigene comprises N-terminal or modified N-

XX CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4

XX CC domains and cysteine-rich domains of dystrophin or utrophin genes. The

XX CC invention also relates to a recombinant adeno-associated virus (AAV)

XX CC comprising dystrophin minigene operably linked to an expression control

QY 1917 CTTCTGCATGATCTCTATCCAAATTCAGACAGAGTTGGGTGAAGTTGCATCCCTTTGGGGC 1976
DB 2980 CTTCTGCATGATCTCTATCCAAATTCAGACAGAGTTGGGTGAAGTTGCATCCCTTTGGGGC 3039
QY 1977 AGTAACATTGAGCCAAAGTGTCCGGA 2001
DB 3040 AGTAACATTGAGCCAAAGTGTCCGGA 3064

RESULT 5
AAD37259
ID AAD37259 standard; DNA; 4476 BP.
XX
AC AAD37259;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-3510.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Homo sapiens.
OS Unidentified.
OS Chimeric.

XX W0200183695-A2.
PN
XX 08-NOV-2001.
PD
XX
XX 27-APR-2001; 2001W0-US013677.
PF
XX
XX 28-APR-2000; 2000US-0200777P.
PR
XX
XX (XIAO/) XIAO X.
PA

XX Xiao X;
XX
XX WPI; 2002-049342/06.
DR
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.

XX Example 1; Page 63-65; 71pp; English.
PS
XX
XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence
XX
XX Sequence 4476 BP; 1252 A; 1096 C; 1127 G; 1001 T; 0 U; 0 Other;

XX
XX Query Match 79.6%; Score 1593; DB 6; Length 4476;
XX Best Local Similarity 87.7%; Pred. No. 0;
XX Matches 1811; Conservative 0; Mismatches 190; Indels 64; Gaps 4;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAACTCGACCGTTTATCAACAGCTTTAGAAGAA 60
DB 1756 GGCAGTTCATTGATGGAGAGTGAAGTAACTCGACCGTTTATCAACAGCTTTAGAAGAA 1815

QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAACACAGAGAGATTTCTAAT 120

DB 1816 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCAAGGAGAGATTTCTAAT 1875
QY 121 GATGTGAAGTGTGGAAGACCCAGTTTCATCTACTCATGAGGGGTACATGATGGATTCACA 180
DB 1876 GATGTGAAGTGTGGAAGACCCAGTTTCATCTACTCATGAGGGGTACATGATGGATTCACA 1935
QY 181 GCCCATCAGGCGCGGTTGGTAAATATCTCAATTTGGAAAGTAAAGCTGATTTGGAAACAGGA 240
DB 1936 GCCCATCAGGCGCGGTTGGTAAATATCTCAATTTGGAAAGTAAAGCTGATTTGGAAACAGGA 1995
QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGA 300
DB 1996 AAATTATCAGAAGATGAAGAAACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGA 2055
QY 301 TGGGAATGCCTCAGGCTAGCTAGCATGGAATAAACAAGCAATTTACATAGAGTTTTAATG 360
DB 2056 TGGGAATGCCTCAGGCTAGCTAGCATGGAATAAACAAGCAATTTACATAGAGTTTTAATG 2115
QY 361 GATCTCCAGAAATC-GAAACTGAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGA 419
DB 2116 GATCTCCAGAAATCAGAAACTGAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGA 2175
QY 420 ACAAGGAAATGGAGAGAGCCCTCTTGGACCTCATCTTTGAAGACCTTAAACACGCCAAGTA 479
DB 2176 ACAAGGAAATGGAGAGAGCCCTCTTGGACCTCATCTTTGAAGACCTTAAACACGCCAAGTA 2235
QY 480 CAACAACTAAAGTCTCTCAAGAAAGATCTAGAAACAAGCAAGTCAAGGTCAATTCCTC 539
DB 2236 CAACAACTAAAGTCTCTCAAGAAAGATCTAGAAACAAGTCAAGGTCAATTCCTC 2295
QY 540 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGCAACATCTCTAGATGACAGAACCCG 659
DB 2296 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGCAACATCTCTAGATGACAGAACCCG 2415
QY 660 TGGGTCTCTTTACAAGACATCCCTCTCAATTTGGCAACGCTTACTGAAGAACAGTGCCTT 719
DB 2416 TGGGTCTCTTTACAAGACATCCCTCTCAATTTGGCAACGCTTACTGAAGAACAGTGCCTT 2475
QY 720 TTTAGTGCATGGCTTTTCAGAAAGAGATGCAGTGAACAAGATTCACACAACCTGGCTTT 779
DB 2476 CTTCTGGTGTGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGGACCTTATTTGGAGGC 2535
QY 780 AAAGATCAAAATGTAATGTTATCAAGTCTTCAAAAACCTGGCGGTTTTAAAAAGCGGATCTA 839
DB 2536 GACTTTCCAGCAGTTTCAGAACAGAACGATGTACATAGGGCCTTCAAGAGGGAATTTGAAA 2595
QY 840 GAAAGAAAAAGCANTCCATGGCAAACTGTA-----TTTCACTCAAAACAG 885
DB 2596 ACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAG 2655
QY 886 ATCTTCTTTCAACACTGAAGAAATAGTCAAGTGAACCCAGAACCGGAAGCATGGCTGGA-- 943
DB 2656 CTTTGGAGGACTAGAGAACTCTTACAGAGGCCAGAGAGCTGCTCTTGGAGAGAGA 2715
QY 944 -----TACTTTGCCGCTGTGGGATTAATTTAGTCCAAA 978
DB 2716 GCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTAGTGGGAAA 2775
QY 979 AACTTTGAA-----AAGAGTACAGCACAGACCCCTTGAAGAAGCTC 1016
DB 2776 AAATTGAACCTGCATCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAGAGCTC 2835
QY 1017 CAGGAACCTTCAAGAGGCCACGGATGAGCTGAGCACTCAAGCTGGCCCAAGCTGAGGTGATC 1076
DB 2836 CAGGAACCTTCAAGAGGCCACGGATGAGCTGAGCACTCAAGCTGGCCCAAGCTGAGGTGATC 2895
QY 1077 AAGGATCTGCGACCGCTGGCGATCTCTCTCAATGACTCTCTCTCAAGATCACTCTGAG 1136
DB 2896 AAGGATCTGCGACCGCTGGCGATCTCTCTCAATGACTCTCTCTCAAGATCACTCTGAG 2955

QY 1137 AAGTCAAGGCACTTCGAGGAGAAATTCGCGCTGTAAGAGAACTGAGCCACGTCAT 1196
DB 2956 AAGTCAAGGCACTTCGAGGAGAAATTCGCGCTGTAAGAGAACTGAGCCACGTCAT 3015
QY 1197 GACCTTGTCCGACGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCACT 1256
DB 3016 GACCTTGTCCGACGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCACT 3075
QY 1257 CTGAAGACCTGACACGACGATGAGCTTCTGAGGTGCGCTGAGGACCGAGTCAGG 1316
DB 3076 CTGAAGACCTGACACGACGATGAGCTTCTGAGGTGCGCTGAGGACCGAGTCAGG 3135
QY 1317 CAGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGACATCTTCTTCCAGGTCT 1376
DB 3136 CAGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGACATCTTCTTCCAGGTCT 3195
QY 1377 GTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCCAC 1436
DB 3196 GTCCAGGTCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCCAC 3255
QY 1437 GAGACTCAAAACAACTTCTGCGGACCATCCCAAATGACAGCTCTACAGTCTTTAGCT 1496
DB 3256 GAGACTCAAAACAACTTCTGCGGACCATCCCAAATGACAGCTCTACAGTCTTTAGCT 3315
QY 1497 GACCTGAATATGTCAGATCTCAGCTTATAGGACTGCCATGAAATCTCCGAAGACTGCAG 1556
DB 3316 GACCTGAATATGTCAGATCTCAGCTTATAGGACTGCCATGAAATCTCCGAAGACTGCAG 3375
QY 1557 AAGGCCCTTTGCTTGGATCTTTCAGCTCTCAGCTGTCATGCTGAGCTGAGCCAGCAC 1616
DB 3376 AAGGCCCTTTGCTTGGATCTTTCAGCTCTCAGCTGTCATGCTGAGCTGAGCCAGCAC 3435
QY 1617 AACCTCAAGCAAAATGACAGCCCATGGATATCTCGAGATTAATTTGTTGACCACT 1676
DB 3436 AACCTCAAGCAAAATGACAGCCCATGGATATCTCGAGATTAATTTGTTGACCACT 3495
QY 1677 ATTATGACGCTGAGGAGAGAGACACAAATTTGGTCAAGCTCCTCTCTCGGTGGAT 1736
DB 3496 ATTATGACGCTGAGGAGAGAGACACAAATTTGGTCAAGCTCCTCTCTCGGTGGAT 3555
QY 1737 ATGTGTCTGAACCTGGCTGCTGAATTTATGATACGGGACGAAACAGGAGGATCGGTGC 1796
DB 3556 ATGTGTCTGAACCTGGCTGCTGAATTTATGATACGGGACGAAACAGGAGGATCGGTGC 3615
QY 1797 CTGTCTTTTAAACTGCGATCATTTCCCTGTGTGTAAGACATTTGGAAGACAAAGTACAGA 1856
DB 3616 CTGTCTTTTAAACTGCGATCATTTCCCTGTGTGTAAGACATTTGGAAGACAAAGTACAGA 3675
QY 1857 TACCTTTTCAAGCAAGTGGCAAGTCAACAGAGATTTTGTACCGAGCGAGGCTGGGCTC 1916
DB 3676 TACCTTTTCAAGCAAGTGGCAAGTCAACAGAGATTTTGTACCGAGCGAGGCTGGGCTC 3735
QY 1917 CTCTGTCATGATCTATFCCAAATTCGAAGACAGTTCGGGTGAAAGTTCATCTTTGGGGC 1976
DB 3736 CTCTGTCATGATCTATFCCAAATTCGAAGACAGTTCGGGTGAAAGTTCATCTTTGGGGC 3795
QY 1977 AGTAAACATTGAGCCAAAGTGTCCCGA 2001
DB 3796 AGTAAACATTGAGCCAAAGTGTCCCGA 3820

RESULT 6

AAZ48568

ID AAZ48568 standard; cDNA to mRNA; 4402 BP.

XX AAZ48568;

XX 31-MAR-2000 (first entry)

XX A rod shortened dystrophin (deltaDysAH3) encoding nucleotide sequence.

XX Muscular dystrophy; rod domain; adeno-associated virus; AAV;

XX

XX

XX

KW dystrophin gene; truncated; ds.

OS Homo sapiens.

XX JP11318467-A.

XX 24-NOV-1999.

XX 08-MAY-1999; 98JP-00142134.

XX 08-MAY-1999; 98JP-00142134.

XX (KAGU-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX (KOKU-) KOKURITSU SEISHIN SHINKEI CENT SOCHO.

XX WPI: 2000-100771/09.

XX P-PSDB; AAY59239.

XX A gene encoding a shortened dystrophin - useful for the treatment of muscular dystrophy.

XX Claim 7; Page 21-22; 44pp; Japanese.

XX The invention provides a gene for the treatment of muscular dystrophy having at least one rod repeat structure of hinge 1, hinge 4 and rod domain of dystrophin gene and having a base sequence of 4.5 kb. The gene and a gene-introducing medium consisting of an adeno-associated virus (AAV) vector or lentivirus vector containing the rod shortened dystrophin genes can be used for the genetic treatment of muscular dystrophy of low immune reaction. The present sequence represents a rod shortened dystrophin encoding sequence

XX Sequence 4402 BP; 1339 A; 984 C; 1010 G; 1069 T; 0 U; 0 Other;

SQ Query Match 66.68; Score 1332; DB 3; Length 4402;

Best Local Similarity 83.88; Pred No. 0;

Matches 1677; Conservative 0; Mismatches 0; Indels 325; Gaps 2;

QY 1 GGCAGTTCATGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 60

DB 1195 GGCAGTTCATGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 1254

QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCAGAGAGATTTCTAAT 120

DB 1255 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCAGAGAGATTTCTAAT 1314

QY 121 GATGTGAAGTGTGTAAGAACCCAGTTCATCTACTCATGAGGGGTACATGATGGATTTGACA 180

DB 1315 GATGTGAAGTGTGTAAGAACCCAGTTCATCTACTCATGAGGGGTACATGATGGATTTGACA 1374

QY 181 GCCCATCAGGCGCGGTTGGTAAATATTTCTACAATTTGGGAAGTAAAGTATGGAACAGGA 240

DB 1375 GCCCATCAGGCGCGGTTGGTAAATATTTCTACAATTTGGGAAGTAAAGTATGGAACAGGA 1434

QY 241 AAATTTATCAGAAGATGAAGAAATGAAGTAAACCTGAAGTAAAGCAGATGAATCTCTTAAATTCAGA 300

DB 1435 AAATTTATCAGAAGATGAAGAAATGAAGTAAACCTGAAGTAAAGCAGATGAATCTCTTAAATTCAGA 1494

QY 301 TGGGAATGCCCTCAGGCTAGCTAGCATGGAAAAAACAAGCAATTTACATAGATTTTAATG 360

DB 1495 TGGGAATGCCCTCAGGCTAGCTAGCATGGAAAAAACAAGCAATTTACATAGATTTTAATG 1554

QY 361 GATCTCCAGAAATC-GAAACTGAAAGAGTTGAAAGACTTGAAGTCTGCTTAAACAAACAGAGAAAGA 419

DB 1555 GATCTCCAGAAATC-GAAACTGAAAGAGTTGAAAGACTTGAAGTCTGCTTAAACAAACAGAGAAAGA 1614

QY 420 ACAAGAAAAATGGAGGAGAGCCCTCTTGGACCTGTATCTTGAAGACCTTAAACGCCCAAGTA 479

DB 1615 ACAAGAAAAATGGAGGAGAGCCCTCTTGGACCTGTATCTTGAAGACCTTAAACGCCCAAGTA 1674

QY 480 CAACAACATAGGTGCTTCAAGAGATCTAGAACAGACACAGTCAAGGTCATTTCTCTC 539

DB 1675 CAACAACATAGGTGCTTCAAGAGATCTAGAACAGACACAGTCAAGGTCATTTCTCTC 1734

QY	540	ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA	599
DB	1735	ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA	1794
QY	600	GAACAACCTTAAGGTATTGGAGATCGATGGGCAACAATCTGTAGATGGACAGAGACCG	659
DB	1795	GAACAACCTTAAGGTATTGGAGATCGATGGGCAACAATCTGTAGATGGACAGAGACCG	1854
QY	660	TGGTTCCTTTTCAAGACATCCTTTCTCAATGGCAACGCTTTACTCTACTCAAGAACAGTGCCTT	719
DB	1855	TGGTTCCTTTTCAAGACATCCTTTCTCAATGGCAACGCTTTACTCTACTCAAGAACAGTGCCTT	1914
QY	720	TTTAGTCATGGCTTTCAGAAAAAGAAGATGCAATGAAACAAGATTCACACAACCTGGCTTT	779
DB	1915	TTTAGTCATGGCTTTCAGAAAAAGAAGATGCAATGAAACAAGATTCACACAACCTGGCTTT	1974
QY	780	AAAGATCAAAATGAAATGTTTATCAAGTCCTTCAAAAACTGGCCGCTTTTAAAGCGGATCTA	839
DB	1975	AAAGATCAAAATGAAATGTTTATCAAGT-----	2001
QY	840	GA AAAAGAAAAGCAATCCATGGGCAAACTGTATTCACTCAACAAGATCTTCTTTCAACA	899
DB	2002	-----	2001
QY	900	CTGAAGAATAAGTCAGTCACCCAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGT	959
DB	2002	-----	2001
QY	960	TGGGATTAATTTAGTCCAAAAAATTTGAAAAAGTACAGCACAGACCCTTGAAGACTCCAG	1019
DB	2002	-----	2001
QY	1020	GAACCTTCAAGAGGCCACGGATAGCTGGACCTCAAGCTGGCGCAAGCTGAGGTGATCAAG	1079
DB	2002	-----	2001
QY	1080	GGATCTCGGACGCCGTGGGGATCTCCTCATTTGACTCTCTCCAGATCACCTCGAGAA	1139
DB	2002	-----	2010
QY	1140	GTCAAGGCACCTTCAGAGAGAAATTCGCGCTCTGAAGAGAAACGTGAGCCACGTCATGAC	1199
DB	2011	GTCAAGGCACCTTCAGAGAGAAATTCGCGCTCTGAAGAGAAACGTGAGCCACGTCATGAC	2070
QY	1200	CTTGCTGCCAGCTTACCATTTTGGGATTCAGCTCTCACGGTATTAACCTCAGGACTCTG	1259
DB	2071	CTTGCTGCCAGCTTACCATTTTGGGATTCAGCTCTCACGGTATTAACCTCAGGACTCTG	2130
QY	1260	GAAGACCTGAAACACAGATGAAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAGGCAG	1319
DB	2131	GAAGACCTGAAACACAGATGAAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAGGCAG	2190
QY	1320	CTGCATGAAGCCACAGGGAATTTGGTCCAGCATCTCAGCACTTTCTTTTCCAGTCTGTG	1379
DB	2191	CTGCATGAAGCCACAGGGAATTTGGTCCAGCATCTCAGCACTTTCTTTTCCAGTCTGTG	2250
QY	1380	CAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACGAG	1439
DB	2251	CAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACGAG	2310
QY	1440	ACTCAAAACAATTCGTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGAC	1499
DB	2311	ACTCAAAACAATTCGTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGAC	2370
QY	1500	CTGAAATATGTCAGATTCCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTCCAGAAG	1559
DB	2371	CTGAAATATGTCAGATTCCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTCCAGAAG	2430
QY	1560	GCCCTTTGCTTGGATCTCTTTGAGCCTGTGACTGATGTATGCTTGGACACGACAAC	1619
DB	2431	GCCCTTTGCTTGGATCTCTTTGAGCCTGTGACTGATGTATGCTTGGACACGACAAC	2490

Qy	1620	CTCAAGCAAAATGACAGCGCCATGGATATCCTGCGAGATTATTAATGTTGTGACCACTATT	1679
Db	2491	CTCAAGCAAAATGACAGCGCCATGGATATCCTGCGAGATTATTAATGTTGTGACCACTATT	2550
Qy	1680	TATGACGGCTGAGCAAGAGCACAAACAATTTGGTCAACGTCCTCTCTGCGTGGATATG	1739
Db	2551	TATGACGGCTGAGCAAGAGCACAAACAATTTGGTCAACGTCCTCTCTGCGTGGATATG	2610
Qy	1740	TGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGAGATCCGTGTCCCTG	1799
Db	2611	TGTCCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGAGATCCGTGTCCCTG	2670
Qy	1800	TCCTTTAAACTGGCATTCATTCCTCTGTGTTAAGCACATTTGGAGACAAAGTACAGATAC	1859
Db	2671	TCCTTTAAACTGGCATTCATTCCTCTGTGTTAAGCACATTTGGAGACAAAGTACAGATAC	2730
Qy	1860	CTTTTCAAGCAAGTGCGCAAGTTCACACAGATTTTGTGACCGCAGCGAGGCTGGGCCCTCCTT	1919
Db	2731	CTTTTCAAGCAAGTGCGCAAGTTCACACAGATTTTGTGACCGCAGCGAGGCTGGGCCCTCCTT	2790
Qy	1920	CTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGAAGTTGCAFTCTCTTTGGGGCAGT	1979
Db	2791	CTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGAAGTTGCAFTCTCTTTGGGGCAGT	2850
Qy	1980	AACATTGAGCCAAAGTGTCCGGA	2001
Db	2851	AACATTGAGCCAAAGTGTCCGGA	2872

RESULT 7

RESULT 7	
AAZ48567	
ID	AAZ48567 standard; cDNA to mRNA; 4402 BP.
XX	
XX	
AC	AAZ48567;
XX	
DT	31-MAR-2000 (first entry)
XX	
DE	A rod shortened dystrophin (deltaDysAx11) encoding nucleotide sequence.
XX	
KW	Muscular dystrophy; rod domain; adeno-associated virus; AAV;
KW	dystrophin gene; truncated; ds.
XX	
OS	Homo sapiens.
XX	
PN	JP11318467-A.

XX	
XX	
PD	24-NOV-1999.
XX	
Pf	08-MAY-1998; 98JP-00142134.
XX	
XX	
PR	08-MAY-1998; 98JP-00142134.
XX	
XX	(KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
PA	(KOKU-) KOKURITSU SEISHIN SHINKEI CENT SOCHO.
FA	
XX	
DR	WPI; 2000-100771/09.
DR	P-ESDB; AAY59238.
XX	
PT	A gene encoding a shortened dystrophin - useful for the treatment of
PT	muscular dystrophy.
XX	
PS	Claim 7; Page 16-17; 44pp; Japanese.
CC	
XX	The invention provides a gene for the treatment of muscular dystrophy
CC	having at least one rod repeat structure of hinge 1, hinge 4 and rod
CC	domain of dystrophin gene and having a base sequence of 4.5 kb. The gen
CC	and a gene-introducing medium consisting of an adeno-associated virus
CC	(AAV) vector or lentivirus vector containing the rod shortened dystroph
CC	genes can be used for the genetic treatment of muscular dystrophy of lo
CC	immune reaction. The present sequence represents a rod shortened
CC	dystrophin encoding sequence
XX	
SQ	Sequence 4402 BP: 1329 A: 1000 C: 1019 G: 1054 T: 0 U: 0 Other;

AAD37237;
 XX DT 21-AUG-2002 (first entry)
 XX DE Human dystrophin minigene delta3849.
 XX XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
 KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
 KW Becker muscular dystrophy; ds.
 XX OS Homo sapiens.
 XX XX W0200183695-A2.
 XX PN 08-NOV-2001.
 XX XX 27-APR-2001; 2001WO-US013677.
 XX PF 28-APR-2000; 2000US-0200777P.
 XX PR (XIAO/) XIAO X.
 XX PA Xiao X;
 XX PI WPI; 2002-049342/06.
 XX DR New dystrophin minigene for treating Duchenne or Becker muscular
 PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
 PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
 PT gene.
 XX PS Example 1; Page 48-49; 71pp; English.
 XX CC The present invention relates to an isolated nucleotide sequence encoding
 CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
 CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
 CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
 CC invention also relates to a recombinant adeno-associated virus (AAV)
 CC comprising dystrophin minigene operably linked to an expression control
 CC element. The dystrophin minigene in operable linkage with an expression
 CC control element, in a recombinant adeno-associated virus or retrovirus is
 CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
 CC dystrophy (BMD) in a mammalian subject. The present sequence is human
 CC dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
 CC hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
 CC and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
 XX SQ Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 U; 0 Other;
 Query Match 63.1%; Score 1263; DB 6; Length 3858;
 Best Local Similarity 75.3%; Pred. No. 0;
 Matches 1816; Conservative 0; Mismatches 185; Indels 412; Gaps 3;
 QY 1 GGCAGTTCATTGATGGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 60
 DB 1000 GGCAGTTCATTGATGGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 1059
 QY 61 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTGCAAGCACCAAGGAGAGATTCTAAT 120
 DB 1060 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTGCAAGCACCAAGGAGAGATTCTAAT 1119
 QY 121 GATGTGGAAGTGGTGAAGACAGTTTCATCTACTCATGAGGGGTACATGATGGATTGACA 180
 DB 1120 GATGTGGAAGTGGTGAAGACAGTTTCATCTACTCATGAGGGGTACATGATGGATTGACA 1179
 QY 181 GCCCATCAGGCGCGGTGGTAAATTTCTCAATTTGGGAAGTAAAGCTGATTGGAAACAGGA 240
 DB 1180 GCCCATCAGGCGCGGTGGTAAATTTCTCAATTTGGGAAGTAAAGCTGATTGGAAACAGGA 1239
 QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGA 300
 DB 1240 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGA 1299

QY 301 TGGGAATGCCCTCAGGCTAGCTAGCATGGAACAAAGAGCAATTTACATAGAGTTTTAATG 360
 DB 1300 TGGGAATGCCCTCAGGCTAGCTAGCATGGAACAAAGAGCAATTTACATAGAGTTTTAATG 1359
 QY 361 GATCTCCAGAAATC-GAAACTGAAAGAGTTGAATGACTGGCTTAAACAAACACAGAGAAAGA 419
 DB 1360 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTTAAACAAACACAGAGAAAGA 1419
 QY 420 ACNAGGAAATGGAGGAGCGCTCTTGGACCTGATCTTGAAGACCTTAAACAGCCCAAGTA 479
 DB 1420 ACNAGGAAATGGAGGAGCGCTCTTGGACCTGATCTTGAAGACCTTAAACAGCCCAAGTA 1479
 QY 480 CAACAACATAAGTGGCTTCAAGAAAGATCTAGAACAGAACAAAGTCAGGTCATTTCTCTC 539
 DB 1480 CAACAACATAAGTGGCTTCAAGAAAGATCTAGAACAGAACAAAGTCAGGTCATTTCTCTC 1539
 QY 540 ACTCATGTGGTGGTGTGATGATCTTAGTGAGATCAGCAACTGCTGCTCTTGGAA 599
 DB 1540 ACTCATGTGGTGGTGTGATGATCTTAGTGAGATCAGCAACTGCTGCTCTTGGAA 1599
 QY 600 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGACAGAGACCGC 659
 DB 1600 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGACAGAGACCGC 1659
 QY 660 TGGGTCTTTTACAAGACATCTTCTCAATGGCAACCTCTTACTGAAGAACAGTGCCTT 719
 DB 1660 TGGGTCTTTTACAAGACATCTTCTCAATGGCAACCTCTTACTGAAGAACAGTGCCTT 1719
 QY 720 TTTAGTGCATGCTTTTCAAGAAAGACAGATGAGTGAACAAAGATTTCACAACTGGCTTT 779
 DB 1720 TTTCTGTGCTGGCTTTACAGAGCTGAAACAACTGCAATGTCTTACAGAGATGCTACCGT 1779
 QY 780 A-----AAGATCAAAATGAAATGTTTATCAAGTCTTTCAAAAACTGGCGTT 824
 DB 1780 AAGGAAAGGCTCCTAGAAAGACTCCAAAGGAGTAAAGAGCTGATGAAACAAATGGCAAGAC 1839
 QY 825 TTAAAGCGGATCTAGAAAGAAAGCAATCCATGGGCAACTCTATTCTACTCAAAACA 884
 DB 1840 CTCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAACCTGGATGAAACAGCCAA 1899
 QY 885 GATCTCTTTTCAACACTGAAGAAATGAATCAAGTGAACCCAGAAAGACGGAAGCATGGCTGAT 944
 DB 1900 AAAATCCTGAGATCCTTGGAGGTTCCGATGATGACAGTCTGTTTACAAAGAGCTTTGGAT 1959
 QY 945 AACTTTGCCCGGTGTGGGATTAATTTAGTCCAAAACCTTGAAGAGATACAGCAC----- 999
 DB 1960 AACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAGGTCCCATTTG 2019
 QY 1000 ----- 999
 DB 2020 GAAGCCAGTTCTGACCAAGTGGAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGG 2079
 QY 1000 ----- 999
 DB 2080 CTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGGACTTTCCAGCA 2139
 QY 1000 ----- 999
 DB 2140 GTTCAGAGCAGAACGATGATCATAGGCGCTTCAAGAGGGAATTTGAAAACCTAAAGAACCT 2199
 QY 1000 ----- 999
 DB 2200 GTAATCATGATGACTTTTGAGACTGTACGAATATTCTGACAGAGCAGCGCTTTGGAGGA 2259
 QY 1000 ----- 999
 DB 2260 CTAGAGAAACTCTACCAGGAGCCGAGAGAGCTGCTCTCTGAGGAGAGAGAGCCAGAGATGTC 2319
 QY 1000 ----- 999
 DB 2320 ACTGGGCTTCTAGAAAGCAGGCTGAGGAGGTCAATCTAGTGGGAAAAATTTGAACCTG 2379
 QY 1000 -----AGACCCCTTGAAGAGACTCCAGGAACCTTCA 1028

Db	2380	CACTCCGCTGACTGCGCAGAGAAAATAGATGAGACCCCTTGAAGAGCTCCAGGAACCTTCAA	2439
Qy	1029	GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTCG	1088
Db	2440	GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCCTCG	2499
Qy	1089	CAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAGGCA	1148
Db	2500	CAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAGTCAGGCA	2559
Qy	1149	CTTCGAGGAGAAATTCGCGCTCTGAAAGAGAAAGTGAGCCACGTCATGACCTTGTCTCGC	1208
Db	2560	CTTCGAGGAGAAATTCGCGCTCTGAAAGAGAAAGTGAGCCACGTCATGACCTTGTCTCGC	2619
Qy	1209	CAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAGACCTG	1268
Db	2620	CAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAGACCTG	2679
Qy	1269	AACACCCAGATGGAAGCTTTCTGCAGGTGGCGGTTCGAGGACCGAGTCAGCGAGCTGCAATGA	1328
Db	2680	AACACCCAGATGGAAGCTTTCTGCAGGTGGCGGTTCGAGGACCGAGTCAGCGAGCTGCAATGA	2739
Qy	1329	GCCACACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGTCAGGGTCCC	1388
Db	2740	GCCACACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGTCAGGGTCCC	2799
Qy	1389	TGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTTACTATTAACACACGAGACTCAAAAC	1448
Db	2800	TGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTTACTATTAACACACGAGACTCAAAAC	2859
Qy	1449	ACTTGTGGGACCATCCCAAATGACAGACTCTACAGCTCTTTAGCTGACCTGGAATAAT	1508
Db	2860	ACTTGTGGGACCATCCCAAATGACAGACTCTACAGCTCTTTAGCTGACCTGGAATAAT	2919
Qy	1509	GTCAGATTTCTCAGCTTATAGACTTGCCATCAAACTCGAGACTGCGAGAGGCGCTTTGC	1568
Db	2920	GTCAGATTTCTCAGCTTATAGACTTGCCATCAAACTCGAGACTGCGAGAGGCGCTTTGC	2979
Qy	1569	TTGATCTCTTGACCTGTGCAGCTGCATGTGATGCTTGGACCGACACCACTCAAGCAA	1628
Db	2980	TTGATCTCTTGACCTGTGCAGCTGCATGTGATGCTTGGACCGACACCACTCAAGCAA	3039
Qy	1629	AATGACCAGCCATGGATTCCTGCAGATTAAATTTGTTGACCACTATTTATGACCCG	1688
Db	3040	AATGACCAGCCATGGATTCCTGCAGATTAAATTTGTTGACCACTATTTATGACCCG	3099
Qy	1689	CTGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTCGGTGGATATGTCTCTGAAC	1748
Db	3100	CTGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTCGGTGGATATGTCTCTGAAC	3159
Qy	1749	TGCGCTGCTGAATTTATGATACGGACGACAGGAGGATCCGTGCTGTCTTTTAAA	1808
Db	3160	TGCGCTGCTGAATTTATGATACGGACGACAGGAGGATCCGTGCTGTCTTTTAAA	3219
Qy	1809	ACTGGCATCATTTCCCTGTGTAAGCACATTTTGGAAAGACAAGTACAGATACCTTTTCAAG	1868
Db	3220	ACTGGCATCATTTCCCTGTGTAAGCACATTTTGGAAAGACAAGTACAGATACCTTTTCAAG	3279
Qy	1869	CAAGTGGCAAGTTCAACAGGATTTTGTGACCGCGAGGCTGGGCCCTCTCTTCGTGATGAT	1928
Db	3280	CAAGTGGCAAGTTCAACAGGATTTTGTGACCGCGAGGCTGGGCCCTCTCTTCGTGATGAT	3339
Qy	1929	TCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCACTCTTTGGGGCAGTAACTGAG	1988
Db	3340	TCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCACTCTTTGGGGCAGTAACTGAG	3399
Qy	1989	CCAAGTGTCCGA	2001
Db	3400	CCAAGTGTCCGA	3412

AAD37257
 ID AAD37257 standard; DNA; 4825 BP.
 XX
 AC AAD37257;
 XX
 DT 21-AUG-2002 (first entry)
 XX
 DE Adeno-associated virus vector plasmid, AAV-MCK-delta3849.
 XX
 KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
 XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
 KW Becker muscular dystrophy; ds.
 XX
 OS Homo sapiens.
 OS Unidentified.
 OS Chimeric.
 XX
 PN WO200183695-A2.
 XX
 PD 08-NOV-2001.
 XX
 PF 27-APR-2001; 2001WO-US013677.
 XX
 PR 28-APR-2000; 2000US-0200777P.
 XX
 PA (XIAO/) XIAO X.
 XX
 PI Xiao X;
 XX
 DR WPI; 2002-049342/06.
 XX
 PT New dystrophin minigene for treating Duchenne or Becker muscular
 PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
 PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
 PT gene.
 XX
 PS Example 1; Page 61-62; 71pp; English.
 XX
 CC The present invention relates to an isolated nucleotide sequence encoding
 CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
 CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
 CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
 CC invention also relates to a recombinant adeno-associated virus (AAV)
 CC comprising dystrophin minigene operably linked to an expression control
 CC element. The dystrophin minigene in operable linkage with an expression
 CC control element, in a recombinant adeno-associated virus or retrovirus is
 CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
 CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
 CC vector plasmid construct containing human dystrophin minigenes, a muscle
 CC creatine kinase (MCK) promoter and a small polyA signal sequence
 XX
 SQ Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 U; 0 Other;
 XX

QY 1000 ----- 999
Db 3040 CTAGAGAAACTCTACAGGAGCCAGAGAGCTGCCTCTGAGGAGAGAGAGAGAGAGATGTC 3099
QY 1000 ----- 999
Db 3100 ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACCTGAGTGGGAAAAATTGAACCTG 3159
QY 1000 ----- 999
Db 3160 CACTCGCTGACTGGGAGAGAAAAATAGATGAGAGCCCTTGAAGACTCCAGGAACCTTCA 3219
QY 1029 GAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGATCCTGG 1088
Db 3220 GAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGATCCTGG 3279
QY 1089 CAGCCGCTGGGCAATCTCTCATTTGACTCTCTCCAGATCACTCTGAGAAAGTCAAGGCA 1148
Db 3280 CAGCCGCTGGGCAATCTCTCATTTGACTCTCTCCAGATCACTCTGAGAAAGTCAAGGCA 3339
QY 1149 CTTTCGAGGAAATTTGGCCTCTGAAAGAGAACTGAGCCACCTCAATGACCTTGTCTGC 1208
Db 3340 CTTTCGAGGAAATTTGGCCTCTGAAAGAGAACTGAGCCACCTCAATGACCTTGTCTGC 3399
QY 1209 CAGCTTACCTTTGGGCAATCTCTCATTTGACTCTCTCCAGATCACTCTGAGAAAGTCAAGGCA 1268
Db 3400 CAGCTTACCTTTGGGCAATCTCTCATTTGACTCTCTCCAGATCACTCTGAGAAAGTCAAGGCA 3459
QY 1269 AACACAGATGGAAGCTTCTGAGGTGGCCGCTGAGGAGCCAGCTCAGGAGCTGCAATGAA 1328
Db 3460 AACACAGATGGAAGCTTCTGAGGTGGCCGCTGAGGAGCCAGCTCAGGAGCTGCAATGAA 3519
QY 1329 GCCCAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTGTCAGGCTGCC 1388
Db 3520 GCCCAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTGTCAGGCTGCC 3579
QY 1389 TGGGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACACAGAGCTCAAA 1448
Db 3580 TGGGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACACAGAGCTCAAA 3639
QY 1449 ACTTGTGGGACCTCCAAATGACAGAGCTTACCAGTCTTTAGTGACCTGAATAAT 1508
Db 3640 ACTTGTGGGACCTCCAAATGACAGAGCTTACCAGTCTTTAGTGACCTGAATAAT 3699
QY 1509 GTGAGATTCAGCTTATAGACTGCGCATGAACTCCGAAAGCTGCGAAGGCGCTTTGC 1568
Db 3700 GTGAGATTCAGCTTATAGACTGCGCATGAACTCCGAAAGCTGCGAAGGCGCTTTGC 3759
QY 1569 TTGATCTCTGAGCTGTGAGCTGCATGTGATGCTTGGACGAGCAACCTCAAGCAA 1628
Db 3760 TTGATCTCTGAGCTGTGAGCTGCATGTGATGCTTGGACGAGCAACCTCAAGCAA 3819
QY 1629 AATGACGAGCCATGATATCTGAGATTTAATTTGTTGACCACTATTTATGACCGC 1688
Db 3820 AATGACGAGCCATGATATCTGAGATTTAATTTGTTGACCACTATTTATGACCGC 3879
QY 1689 CTGAGGAGAGAGCAACAAATTTGGTCAAGCTCCCTCTCTGCGTGGATATGTGCTGAAC 1748
Db 3880 CTGAGGAGAGAGCAACAAATTTGGTCAAGCTCCCTCTCTGCGTGGATATGTGCTGAAC 3839
QY 1749 TGGCTGCTGAATGTTATGATACGGGAGCAAGGAGGATCCGCTGCTCTTTTAA 1808
Db 3940 TGGCTGCTGAATGTTATGATACGGGAGCAAGGAGGATCCGCTGCTCTTTTAA 3999
QY 1809 ACTGATCATTTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1868
Db 4000 ACTGATCATTTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4059
QY 1869 CAGGTGGCAAGTTCAACAGGATTTTGTGACAGGAGGCTGGGCTCTCTCTCTCATGAT 1928
Db 4060 CAGGTGGCAAGTTCAACAGGATTTTGTGACAGGAGGCTGGGCTCTCTCTCTCATGAT 4119
QY 1929 TCTATCCAAATCCAAAGACAGTGGGTGAAGTTGCATCTTTTGGGGGAGTAAACATTGAG 1988

Db 4120 TCTATCCAAATCCAAAGACAGTGGGTGAAGTTGCATCTTTGGGGCAGTAAACATTGAG 4179
QY 1989 CCAAGTGTCCGGA 2001
Db 4180 CCAAGTGTCCGGA 4192
RESULT 11
AAD37264
ID AAD37264 standard; DNA; 5060 BP.
XX AAD37264;
AC AAD37264;
XX 21-AUG-2002 (first entry)
XX Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
XX Becker muscular dystrophy; ds.
XX Homo sapiens.
XX Cytomegalovirus.
XX Unidentified.
XX Chimeric.
XX WO200183695-A2.
XX 08-NOV-2001.
XX 27-APR-2001; 2001WO-US013677.
XX 28-APR-2000; 2000US-0200777P.
XX (XIAO/) XIAO X.
XX Xiao X;
XX WPI; 2002-049342/06.
XX New dystrophin minigene for treating Duchenne or Becker muscular
XX dystrophy comprises an N-terminal domain or modified N-terminal domain,
XX rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
XX gene.
XX Example 1; Page 70-71; 7lpp; English.
XX The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified N-
XX terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)
XX comprising dystrophin minigene operably linked to an expression control
XX element. The dystrophin minigene in operable linkage with an expression
XX control element, in a recombinant adeno-associated virus or retrovirus is
XX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX dystrophy (BMD) in a mammalian subject. The present sequence is AAV
XX vector plasmid construct containing human dystrophin minigenes, a muscle
XX creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter and a
XX small polya signal sequence
XX Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 U; 0 Other;
SQ
Query Match 63.1%; Score 1263; DB 6; Length 5060;
Best Local Similarity 75.3%; Pred. No. 0;
Matches 1816; Conservative 0; Mismatches 185; Indels 412; Gaps 3;
QY 1 GGCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAACACACTTTAGAGAA 60
Db 1992 GGCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAACACACTTTAGAGAA 2051
QY 61 GTATTATCGTGGCTTCTTTCGTGAGGACACATTGCAAGCACAGAGAGATTTCTAAT 120

Db 2052 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGGANTTCTAAT 2111
QY 121 GATGTGGAAGTGGTGAAGACAGGTTTCATATCTCATGAGGGGTACATGATGATTTGACA 180
Db 2112 GATGTGGAAGTGGTGAAGACAGGTTTCATATCTCATGAGGGGTACATGATGATTTGACA 2171
QY 181 GCCCATCAGGCCCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGA 240
Db 2172 GCCCATCAGGCCCGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTTGGAACAGGA 2231
QY 241 AAATTTATCAGAAGATGAAGAAGTGAAGTACAAGACAGATGAATCTCTTAATTTCAAGA 300
Db 2232 AAATTTATCAGAAGATGAAGAAGTGAAGTACAAGACAGATGAATCTCTTAATTTCAAGA 2291
QY 301 TGGGATGCTCCTCAGGTAGCTAGCATGGGAAAAAACAAGCAATTTACATAGATTTTAATG 360
Db 2292 TGGGATGCTCCTCAGGTAGCTAGCATGGGAAAAAACAAGCAATTTACATAGATTTTAATG 2351
QY 361 GATCTCCAGAATC-GAAACTGAAAGAGTTGAATGACTTGAAGACCTTAAACCCCAAGTA 419
Db 2352 GATCTCCAGAATCAGAATCAGAAGTTGAATGACTTGAAGACCTTAAACCCCAAGTA 2411
QY 420 ACAGGAAAAATGGAAGGAGAGCTCTTTGGACCTGATCTTGAAGACCTTAAACCCCAAGTA 479
Db 2412 ACAGGAAAAATGGAAGGAGAGCTCTTTGGACCTGATCTTGAAGACCTTAAACCCCAAGTA 2471
QY 480 CAACAACATAGGTGCTTCAAGAGATCTAGAGACCAAGCAAGTCAAGGTCAATTTCTCTC 539
Db 2472 CAACAACATAGGTGCTTCAAGAGATCTAGAGACCAAGCAAGTCAAGGTCAATTTCTCTC 2531
QY 540 ACTCACATGTTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 599
Db 2532 ACTCACATGTTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2591
QY 600 GAACAACATTAAGTTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 659
Db 2592 GAACAACATTAAGTTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 2651
QY 660 TGGTCTCTTTTACAAGACATCTCTCAAAATGCAACGCTTACTTGAAGAACAGTGCCTT 719
Db 2652 TGGTCTCTTTTACAAGACATCTCATAGATTACTGCAACAGTTCCTCCCTGACCTTGGAAAG 2711
QY 720 TTATGTCATGGCTTTCAGAAAAAGAGATGCGAGTGAACAAAGATTACACAACATGGCTTT 779
Db 2712 TTTCTTGGCTTACAGAACTGTAACCAACTGCAATGCTCCTACAGGATGCTACCCGT 2771
QY 780 A-----AAGATCAAAATGAATGTTATCAAGTCTTCAAAACCTGGCCGT 824
Db 2772 AAGAAAGGCTCTAGAAAGCTCCAGGGAGTAAAGAGCTGATGAACAATGGCAAGAC 2831
QY 825 TTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCATGGGCAAACTGATTTCACTCAAAACAA 884
Db 2832 CTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAACACTGGATGAACACAGCCAA 2891
QY 885 GATCTCTTTTCAACATGAGAAATAGTCAGTGACCCAGAAAGCAAGCATGGCTGGAT 944
Db 2892 AAATCTGAGTCCCTTGGAGGTTCCGATGTCAGTCCCTGTTTCAAAAGAGCTTTGGAT 2951
QY 945 AACTTTGCCCGGTGTTGGGATAATTAGTCCAAAAACCTTGAAGAGAGTACAGAC----- 999
Db 2952 AACATGAACCTTCAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACATTAGTCCCATTTG 3011
QY 1000 ----- 999
Db 3012 GAAGCCAGTTCTGACCAAGTGGAGCGTCTGCACCTTTCTCTGAGGAACCTTCTGGTGG 3071
QY 1000 ----- 999
Db 3072 CTACAGCTGAAGATGATGAATTAAGCCGAGGACCTATTGGAGGGACCTTTCCAGCA 3131
QY 1000 ----- 999

Db 3132 GTTCAGAACGACGATGTACATAGGCGCTTCAAGAGGGAAATTGAAAACTAAAGAACCT 3191
QY 1000 ----- 999
Db 3192 GTAATCATGACTCTTTGAGACTGTACGAATATTTCTGACAGACAGCCTTTTGAAGGA 3251
QY 1000 ----- 999
Db 3252 CTAGAGAAACTCTACCAGGAGCCAGAGAGTGCCTCTCTGAGGAGAGAGCCAGAAATGC 3311
QY 1000 ----- 999
Db 3312 ACTCGCTTCTAGGAAAGCAGGCTGAGGAGTCAATATCTGAGTGGGAAAAAATTGAACCTG 3371
QY 1000 -----AGACCTTTGAAAGACTCCAGGAACTTCAA 1028
Db 3372 CACTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAA 3431
QY 1029 GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGTGATCAAGGGATCTCTGG 1088
Db 3432 GAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGTGATCAAGGGATCTCTGG 3491
QY 1089 CAGCCCGTGGGCGATCTCTCATTTGACTCTCTCCTCAAGATCACCTCGAGAAAAAGTCAAGCA 1148
Db 3492 CAGCCCGTGGGCGATCTCTCATTTGACTCTCTCCTCAAGATCACCTCGAGAAAAAGTCAAGCA 3551
QY 1149 CTTGAGGAGAAAAATGCGCTCTGAAAGAGAAAGTGAGCCACGCTCAATGACTCTTCTGCTGC 1208
Db 3552 CTTGAGGAGAAAAATGCGCTCTGAAAGAGAAAGTGAGCCACGCTCAATGACTCTTCTGCTGC 3611
QY 1209 CAGCTTACACTTTTGGGCACTTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 1268
Db 3612 CAGCTTACACTTTTGGGCACTTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 3671
QY 1269 AACACAGATGGAAGCTTTGCAAGTGGCGTCAAGGACCGAGTCAAGGAGCTGCAATGAA 1328
Db 3672 AACACAGATGGAAGCTTTGCAAGTGGCGTCAAGGACCGAGTCAAGGAGCTGCAATGAA 3731
QY 1329 GCCCACAGGACTTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGCTGTCAGGCTCC 1388
Db 3732 GCCCACAGGACTTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGCTGTCAGGCTCC 3791
QY 1389 TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACACAGAGACTCAAAACA 1448
Db 3792 TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACACAGAGACTCAAAACA 3851
QY 1449 ACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACAGCTTTTAGTGAACCTGATTAAT 1508
Db 3852 ACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACAGCTTTTAGTGAACCTGATTAAT 3911
QY 1509 GTCAGATTTCTCAGCTTATAGGACTGCGCATGAAACTCCGAAAGCTGCAAGAGGCGCTTTGC 1568
Db 3912 GTCAGATTTCTCAGCTTATAGGACTGCGCATGAAACTCCGAAAGCTGCAAGAGGCGCTTTGC 3971
QY 1569 TTGATCTCTTGGACCTGTGATGCTGATGCTGCTTGGACCGAGCAACACCTCAAGCAA 1628
Db 3972 TTGATCTCTTGGACCTGTGATGCTGATGCTGCTTGGACCGAGCAACACCTCAAGCAA 4031
QY 1629 AATCAGCAGCCCATGGGATATCTCTGAGATTAATTTGTTTGGACCACTATTTATGACCCG 1688
Db 4032 AATCAGCAGCCCATGGGATATCTCTGAGATTAATTTGTTTGGACCACTATTTATGACCCG 4091
QY 1689 CTGAGCAAGAGCACAACAAATTTGGTCAAGCTCCCTCTCTGCTGGATATGCTCTGAAC 1748
Db 4092 CTGAGCAAGAGCACAACAAATTTGGTCAAGCTCCCTCTCTGCTGGATATGCTCTGAAC 4151
QY 1749 TGGCTGCTGAATGTTTATGATAGCGGACCAAGGAGGATCCGCTCTCTGCTGCTTTTAA 1808
Db 4152 TGGCTGCTGAATGTTTATGATAGCGGACCAAGGAGGATCCGCTCTCTGCTGCTTTTAA 4211
QY 1809 ACTGGCATATTTCCCTGTGTAAGAGCACAATTTGGAAGACAAGTACAGATACCTTTTCAAG 1868
Db 4212 ACTGGCATATTTCCCTGTGTAAGAGCACAATTTGGAAGACAAGTACAGATACCTTTTCAAG 4271

QY 1869 CAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGCTGGCCCTCCTTCTGCATGAT 1928
DB 4272 CAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGCTGGCCCTCCTTCTGCATGAT 4331
QY 1928 TCTATCCAAATTCACAGACAGTTGGTGAAGTTGCATCCTTTGGGGGCGAGTACATGAG 1988
DB 4332 TCTATCCAAATTCACAGACAGTTGGTGAAGTTGCATCCTTTGGGGGCGAGTACATGAG 4391
QY 1989 CCAAGTGTCCGGA 2001
DB 4392 CCAAGTGTCCGGA 4404

RESULT 12

ID ABK81998 standard; DNA; 5339 BP.

XX ABK81998;

DT 13-AUG-2002 (first entry)

XX DNA encoding mini-dystrophin protein deltaR2-R21.

XX Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;

XW Duchenne's muscular dystrophy; DMD; dystrophin; ds.

XX Homo sapiens.

OS Synthetic.

XX WO200229056-A2.

PN 11-APR-2002.

XX 04-OCT-2001; 2001WO-US031126.

XX 06-OCT-2000; 2000US-0238848P.

XX (UNMI) UNIV MICHIGAN.

XX Chamberlain JS, Harper SQ;

PI WPI; 2002-435334/46.

XX A composition for preparing therapeutic drugs, has a mini-dystrophin

PT peptide comprising a specific number of spectrin-like repeat domains, or

PT a nucleic acid sequence encoding the mini-dystrophin peptide.

XX Example 6; Fig 13; 145pp; English.

XX The invention describes a composition comprising a mini-dystrophin

CC peptide comprising a spectrin-like repeat domain, where the domain

CC comprises n spectrin-like repeats, and contains no more than n spectrin-

CC like repeats, where n is an even number between 4-24, or a nucleic acid

CC encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the

CC polynucleotide encoding it is useful as a medicament, for preparing a

CC drug for therapeutic application and in the preparation of a composition

CC for treatment of muscle disease, e.g. Duchenne's muscular dystrophy

CC (DMD). This sequence represents a mini-dystrophin sequence of the

CC invention

XX Sequence 5339 BP; 1638 A; 1191 C; 1187 G; 1323 T; 0 U; 0 Other;

QY Query Match 62.7%; Score 1254; DB 6; Length 5339;

XX Best Local Similarity 77.9%; Pred. No. 0;

XX Matches 1621; Conservative 0; Mismatches 380; Indels 79; Gaps 6;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 60

DB 1199 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 1258

QY 61 GTATTATCGTGGCTTTCTTCTGCTGAGGACATTTGAAGCACAAGGAGAGATTTCTAAT 120

DB 1259 GTATTATCGTGGCTTTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 1318
QY 121 CATGTGGAAGTGGTGAAGACCACTTTTCATCTCATGAGGGGTACATGATGATTTGACA 180
DB 1319 GATGTGGAAGTGGTGAAGACCACTTTTCATCTCATGAGGGGTACATGATGATTTGACA 1378
QY 181 GCCCATCAGSGCCGGGTGGTAAATATCTCAATTTGGGAGTAAAGCTGATTTGGAAACAGGA 240
DB 1379 GCCCATCAGSGCCGGGTGGTAAATATCTCAATTTGGGAGTAAAGCTGATTTGGAAACAGGA 1438
QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTCAGA 300
DB 1439 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCTCTAAATTCAGA 1498
QY 301 TGGGAATCCCTCAGGTTAGCTAGCATGGAAAAACAAGCAATTTTACATAGAGTTTAAATG 360
DB 1499 TGGGAATCCCTCAGGTTAGCTAGCATGGAAAAACAAGCAATTTTACATAGATTTACTG 1558
QY 361 GATCTCCAGAATC-GAAACTCAAAAGATTGAATGACTGGCTAAACAAAAACAAGAAAGA 419
DB 1559 CAACAGTTCCCTGGACCTGGAAAAAGTTTCTTGGCTTACAGAAGCTGAACAACACT 1618
QY 420 ACAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTA 479
DB 1619 GCCAATGTCTACAGGATGCTACCCGTAAAGGAAAGGCTCCTTAGAAGACTCCAAGGGAGTA 1678
QY 480 CAACCAACATAAGTGTCTTCAAGAGATCTAGAACAAGCAACAGTCAGGCTCAATTTCTCTC 539
DB 1679 AAAGAGCTGATGAACAACAAATGGCAAGACCTCCAGAGTGAAATTTGAAGTCAACAGATGT 1738
QY 540 ACTCATGTGGTGGTGGTAG-----TTGATGAATCTAGTGGAGATCAACCAACT 587
DB 1739 TATCAACACCTGGATGAACACAGCCAAAAATCCTGAGATCCCTGGAAGTTCCGATGAT 1798
QY 588 GCTGCTTTGGAAGAACAACTTAAGG---TATTTGGAGATCGATGGGCAACATCTGTAGA 644
DB 1799 GCAGTCTCTTACAAGAGCGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTCCGAAA 1858
QY 645 TGGACAGAAGACCGCTGGGTCTTTTACAAGACATCTTCTCAAAATGGCAACGCTTACT 704
DB 1859 AAGTCTCTCAACATTAGTTCCTTTTGAAGCCAGTCTGACCAGTGAAGCGTCTGCAC 1918
QY 705 GAAGAACAGTGCCTTTTATGTCATGGCTTTCAGAAAAAGAGATGTCAGTGAACAAGATT 764
DB 1919 CTTTCTCTCAGGAACTTCTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGAG 1978
QY 765 CACACAACTGGCTTTAAAGATCAAAATGAATTTTATCAAGTCTTCAAAAATGGCCCTT 824
DB 1979 GCACCTATTGGAGCGACTTTCACAGCTTCAGAAGCAGAGATGTACATAGGCGCTTC 2038
QY 825 TTAAAGCGGATCTAGAAAAAGAAAGCAATCCATGGGCAAACTGTA----- 871
DB 2039 AAGAGGGAATTGAAACATAAGAACCTGTAAATCATGACTACTCTTGAGACTGTACGAATA 2098
QY 872 -TTCACTCAACAAGATCTTTTCAACACTGAAGAAATAGTCACTGACGCCAGAGAGCGG 930
DB 2099 TTTCTGACAGACGACGCTTTTGAAGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTG 2158
QY 931 AAGCATGGCTGA-----TAACTTTTCCCGGCTGTGGG 963
DB 2159 CTTCTGAGGAGAGAGCCAGATGTCTACTGGCTTCTACGAAAGCAGCTGAGGAGTTC 2218
QY 964 ATAATTAGTCCAAAAACTTGAA-----AAGATCAGCAGAG 1001
DB 2219 AATACTAGTGGGAAAAATTTGAACCTCGCTGACTGGCAGAGAAAAATAGATGAG 2278
QY 1002 ACCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGAGACCTCAAGCTGGCC 1061
DB 2279 ACCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGAGACCTCAAGCTGGCC 2338
QY 1062 CAAGCTGAGGTGATCAAGGATTCCTGGAGCCCGTGGGCGATCTCTCTATTGACTCTCTC 1121
DB 2339 CAAGCTGAGGTGATCAAGGATTCCTGGAGCCCGTGGGCGATCTCTCTATTGACTCTCTC 2398

QY 1122 CAAGATCACCTCGAAGAGTCAAGGCACCTTCGAGGAGAAATTTGGCCTCTCTCAAGAGAAC 1181
Db CAAGATCACCTCGAAGAGTCAAGGCACCTTCGAGGAGAAATTTGGCCTCTCTCAAGAGAAC 2458
QY 1182 GTGAGCCAGTCAATGACCTTGTCTGCGAGCTTACCACTTTGGCATTGAGCTCTCACCG 1241
Db GTGAGCCAGTCAATGACCTTGTCTGCGAGCTTACCACTTTGGCATTGAGCTCTCACCG 2518
QY 1242 TATAACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTTCTGACAGTGGCGTC 1301
Db TATAACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGCTTTCTGACAGTGGCGTC 2578
QY 1302 GAGGACGAGTCAGGACGTGATGAAGCCACACAGGACCTTTGGTCCAGCATCTCAGCAC 1361
Db GAGGACGAGTCAGGACGTGATGAAGCCACACAGGACCTTTGGTCCAGCATCTCAGCAC 2638
QY 1362 TTTCTTTCCAGCTGTCTCCAGGGTCCCTGGAGAGACCCATCTCGCCAAACAAAGTGCCC 1421
Db TTTCTTTCCAGCTGTCTCCAGGGTCCCTGGAGAGACCCATCTCGCCAAACAAAGTGCCC 2698
QY 1422 TACTATATCAACACGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTC 1481
Db TACTATATCAACACGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTC 2758
QY 1482 TACCAAGCTTTAGCTGACCTGAATATGTCAGATCTCAGATCTTAGGACTGCATGAA 1541
Db TACCAAGCTTTAGCTGACCTGAATATGTCAGATCTCAGATCTTAGGACTGCATGAA 2818
QY 1542 CTCGGAAGACTGCAAGAGGCCCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGATGAT 1601
Db CTCGGAAGACTGCAAGAGGCCCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGATGAT 2878
QY 1602 GCCTTGGACGACGACCACTCAAGCAAAATGACAGCCCATGATATCTGCAAGATATT 1661
Db GCCTTGGACGACGACCACTCAAGCAAAATGACAGCCCATGATATCTGCAAGATATT 2938
QY 1662 AATTGTTTGACCACTATTATGACCGCTGAGCAGAGACACAAATTTGGTCAACGTC 1721
Db AATTGTTTGACCACTATTATGACCGCTGAGCAGAGACACAAATTTGGTCAACGTC 2998
QY 1722 CCTCTCTGCTGGATATGCTGTGAATGCTGCTGCTGAATGTTTATGATCGGACGAA 1781
Db CCTCTCTGCTGGATATGCTGTGAATGCTGCTGCTGAATGTTTATGATCGGACGAA 3058
QY 1782 GGGAGGATCCGTGCTGCTTTTAAACCTGGCATCATTTCCCTGTGTAAGACACATTTG 1841
Db GGGAGGATCCGTGCTGCTTTTAAACCTGGCATCATTTCCCTGTGTAAGACACATTTG 3118
QY 1842 GAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAG 1901
Db GAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAG 3178
QY 1902 CGCAGGCTGGCCCTCTCTTCTGATGATCTATCCAAATTCGAAGACAGTTGGTGAAGTT 1961
Db CGCAGGCTGGCCCTCTCTTCTGATGATCTATCCAAATTCGAAGACAGTTGGTGAAGTT 3238
QY 1962 GCATCTCTTTGGGGCAGTAACATTGAGCCAAAGTGTCCGGA 2001
Db GCATCTCTTTGGGGCAGTAACATTGAGCCAAAGTGTCCGGA 3278

RESULT 13

AAD37238

ID AAD37238 standard; DNA; 3531 BP.

XX

AC AAD37238;

XX

DT 21-AUG-2002 (first entry)

XX

DE Human dystrophin minigene delta3531.

XX

KW Human dystrophin minigene: muscular: gene therapy: utrophin: spectrin

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
OS Homo sapiens.
XX WO200183695-A2.
PN 08-NOV-2001.
PD 27-APR-2001; 2001WO-US013677.
XX 28-APR-2000; 2000US-0200777P.
PR (XIAO/) XIAO X.
PA Xiao X;
XX WPI; 2002-049342/06.
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX Example 1; Page 50-51; 71pp; English.
PS The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus,
CC hinge H1 and rod R1), 8059-10227 (rods R22, R23 and R24, hinge H4 and CR
CC domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 U; 0 Other;

Query Match 62.5%; Score 1251.2; DB 6; Length 3531;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1623; Conservative 0; Mismatches 378; Indels 85; Gaps 6;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGCTTATCAACAGCTTTAGAGAA 60
Db 1000 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGCTTATCAACAGCTTTAGAGAA 1059
QY 61 GTATTATCTGCTGCTCTTTCTGCTGAGGACACATTGCAAGCAAGAGAGATTTCTTAAT 120
Db 1060 GTATTATCTGCTGCTCTTTCTGCTGAGGACACATTGCAAGCAAGAGAGATTTCTTAAT 1119
QY 121 GATGTGAAGTGGTGAAGAACCCAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 180
Db 1120 GATGTGAAGTGGTGAAGAACCCAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 1179
QY 181 GCCCATCAGGGCCGGGTGGTGAATATCTCAATTTGGAGAGTGAAGCTGATTGACAGGA 240
Db 1180 GCCCATCAGGGCCGGGTGGTGAATATCTCAATTTGGAGAGTGAAGCTGATTGACAGGA 1239
QY 241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGACAGATGAATCTCTCTAAATTTCAAGA 300
Db 1240 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGACAGATGAATCTCTCTAAATTTCAAGA 1299
QY 301 TGGGAATGCCCTCAGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 360
Db 1300 TGGGAATGCCCTCAGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 1359
QY 361 GATCTCCAGAA-----TCGAAACTGAAAGAGTTGAATGACTGGCTTAAACAAAAACAGAA 413
Db 1360 TTAATCTCAACACTTCCCTCCCTGACCTGCAAAAATTTCTTCTCTCTGCTTACAGAGCTGAA 1419

PT New dystrophin minigene for treating Duchenne or Becker muscular
 PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
 PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
 PT gene.

XX Example 1; Page 62-63; 71bp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
 CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
 CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
 CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
 CC invention also relates to a recombinant adeno-associated virus (AAV)
 CC comprising dystrophin minigene operably linked to an expression control
 CC element. The dystrophin minigene in operable linkage with an expression
 CC control element, in a recombinant adeno-associated virus or retrovirus is
 CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
 CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
 CC vector plasmid construct containing human dystrophin minigenes, a muscle
 CC creatine kinase (MCK) promoter and a small polyA signal sequence

XX Sequence 4498 BP; 1251 A; 1118 C; 1123 G; 1006 T; 0 U; 0 Other;

Query Match 62.5%; Score 1251.2; DB 6; Length 4498;

Best Local Similarity 77.8%; Pred. No. 0;
 Matches 1623; Conservative 0; Mismatches 378; Indels 85; Gaps 6;

Qy	1	GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTTATCAACACAGCTTTAGAAGAA	60
Db	1757	GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTTATCAACACAGCTTTAGAAGAA	1816
Qy	61	GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAGCACAAGGAGAGATTTCCTAAT	120
Db	1817	GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAGCACAAGGAGAGATTTCCTAAT	1876
Qy	121	GATGTGAAGTGGTGAAGACCACTTCTACTCATGAGGGGTACATGATGATTTGACA	180
Db	1877	GATGTGAAGTGGTGAAGACCACTTCTACTCATGAGGGGTACATGATGATTTGACA	1936
Qy	181	GCCCATCAGCGCGGGTGTGTAATATTTCTACAATTTGGGAAGTAAAGCTGATTTGAACAGGA	240
Db	1937	GCCCATCAGCGCGGGTGTGTAATATTTCTACAATTTGGGAAGTAAAGCTGATTTGAACAGGA	1996
Qy	241	AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCTTAATTTCAAGA	300
Db	1997	AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCTTAATTTCAAGA	2056
Qy	301	TGGGAATGCTTCAGGTPAGTACATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG	360
Db	2057	TGGGAATGCTTCAGGTPAGTACATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG	2116
Qy	361	GATCTCCAGAA-----TCGAACCTGAAGAGTTGAATGACCTGCTTAATAATCAAGA	413
Db	2117	TACTGCAACAGTTTCCCTCGACCTGGAAAGTTTCTTGGCTGGCTTACAGAAGCTGAA	2176
Qy	414	GAAGAAACAAGGAAATGGAGGAGAGACCTCTTGGACCTGATCTTGAAGACCTTAAACGC	473
Db	2177	ACAACTGCCAATGCTCTACAGGATGCTACCGTGAAGAAAGGCTCCTAGAAGACTCCAG	2236
Qy	474	CAGTACAACAATAGGTGCTTCAAGAGATCTAGAACAGAAACAAGTCAAGGTCAAT	533
Db	2237	GGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACA	2296
Qy	534	TCCTCTCACTCACATGTTGGTGGTGTGATGA-----ATCTAGTGGAGATCAC	581
Db	2297	GATGTTTATCAACACTGGATGAACAGCCCAAAATCTTGATCTCCCTGGAAGTTCC	2356
Qy	582	GCAACTGCTCTTTGGAAGAACACTTAAG-----TATTGGAGATCGATGGCAACAATPC	638
Db	2357	GATGATGCAGTCTCTGTTTACAAGACCGTTTGGATAACATGAATTTCAAGTGGAGTGAAT	2416
Qy	639	TGTAGATGGACAGAACCGCTGGTCTTTTACAAGACATCTTCTCAAAATGGCAAGCT	698
Db	2417	CGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTCTTGACCAAGTGGAGCGT	2476

Qy	699	CTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTTCAGAAAAAGAGATGAGTGAAC	758
Db	2477	CTGCACCTTTCTCTGCAGAACTTCTGTGTGGCTACAGCTGAAGATGATGAATTAAGC	2536
Qy	759	AAGATTTCACAACTGGCTTTTAAAGATCAAAATGTTTATCAAGTCTTCTTAAAAAATG	818
Db	2537	CGGAGGACACTATTGGAGGCGACTTTCCAGCAGTTCAGAAAGCAGACGATGATACATAGG	2596
Qy	819	GCCGTTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCATGGGCAACTGTA-----	871
Db	2597	GCCTTCAAGAGGGAATTGAAACCTTAAAGAACCTGTAATCATGAGTACTTCTTGAGACTGT	2656
Qy	872	-----TTCACTCAAAACAAGATCTTTTCAACACTGAAGAAATAAGTCAAGTGAACCCAGA	924
Db	2657	CGAATATTCTGCACAGACGACCTTTTGAAGGACTAGAGAACTCTACAGGAGCCCGAGA	2716
Qy	925	AGACGAAGCATGGCTGGA-----TAACTTTGCCCGGT	957
Db	2717	GAGCTGCCTCTGAGGAGAGAGCCAGATGTCTACTCGGCTTCTACGAAAGCAGGCTGAG	2776
Qy	958	GTTGGGATAATTAGTCCAAAACTTGAA-----AAGAGTACA	995
Db	2777	GAGTCAATCTGAGTGGGAAAAATTTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATA	2836
Qy	996	GCACAGACCTTTGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAG	1055
Db	2837	GATGAGACCTTTGAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAG	2896
Qy	1056	CTGCGCAAGCTGAGGTGATCAAGGATCTGCGCAGCCGCTGGCGGATCTCTCATTTGAC	1115
Db	2897	CTGCGCAAGCTGAGGTGATCAAGGATCTGCGCAGCCGCTGGCGGATCTCTCATTTGAC	2956
Qy	1116	TCTCTCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCTCTGAAA	1175
Db	2957	TCTCTCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCTCTGAAA	3016
Qy	1176	GAGAACTGAGCCACGTCATGACCTTGTCTGCGCAGCTTACCACTTTTGGGCATTCAGCTC	1235
Db	3017	GAGAACTGAGCCACGTCATGACCTTGTCTGCGCAGCTTACCACTTTTGGGCATTCAGCTC	3076
Qy	1236	TCACGCTATTAACCTCAGCCTCTGGAAGACTGAAACACCAAGATGGAAGCTTCTGCAAGTG	1295
Db	3077	TCACGCTATTAACCTCAGCCTCTGGAAGACTGAAACACCAAGATGGAAGCTTCTGCAAGTG	3136
Qy	1296	GCCGTCAGAGACCGAGTCAGGACGTCATGAAGCCACAGGACCTTTGGTCAGAGATCT	1355
Db	3137	GCCGTCAGAGACCGAGTCAGGACGTCATGAAGCCACAGGACCTTTGGTCAGAGATCT	3196
Qy	1356	CAGCAGCTTTCTTTCCAGCTGTCTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAA	1415
Db	3197	CAGCAGCTTTCTTTCCAGCTGTCTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAA	3256
Qy	1416	GTGCGCTACTATATCAACACGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACA	1475
Db	3257	GTGCGCTACTATATCAACACGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACA	3316
Qy	1476	GAGCTCTACAGCTTTTGTAGCTGACCTGAATTAATGTCAGATTTCTCAGCTTATAGACTGCC	1535
Db	3317	GAGCTCTACAGCTTTTGTAGCTGACCTGAATTAATGTCAGATTTCTCAGCTTATAGACTGCC	3376
Qy	1536	ATGAAACTCCGAGACTTGCAGAGGCGCTTTGTTGGATCTCTTGGACCTGTGAGCTGCA	1595
Db	3377	ATGAAACTCCGAGACTTGCAGAGGCGCTTTGTTGGATCTCTTGGACCTGTGAGCTGCA	3436
Qy	1596	TGTGATGCTTGGACGAGCAACCTCAAGCAAAATGACAGGCCCATGATATCTTCGCAG	1655
Db	3437	TGTGATGCTTGGACGAGCAACCTCAAGCAAAATGACAGGCCCATGATATCTTCGCAG	3496
Qy	1656	ATTATTAAATGTTGACCACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTC	1715
Db	3497	ATTATTAAATGTTGACCACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTC	3556

Qy	1716	AAAGTCCCTCTCTGCGTGGATATGTGCTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGA	1775
Db	3557	AAAGTCCCTCTCTGCGTGGATATGTGCTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGA	3616
Qy	1776	CGAACAGGAGGATCCGTGCTCTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCA	1835
Db	3617	CGAACAGGAGGATCCGTGCTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCA	3676
Qy	1836	CATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAAACAGATTTTGT	1895
Db	3677	CATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAAACAGATTTTGT	3736
Qy	1896	GACCAGCGCAGGCTGGCGCTCCTTCTGCATGATTCTATCAAATTCCAAAGACAGTTGGGT	1955
Db	3737	GACCAGCGCAGGCTGGCGCTCCTTCTGCATGATTCTATCAAATTCCAAAGACAGTTGGGT	3796
Qy	1956	GAAGTTGCATCCTTTTGGGGGCAGTAACATTGAGCCAAAGTGTCCGGA	2001
Db	3797	GAAGTTGCATCCTTTTGGGGGCAGTAACATTGAGCCAAAGTGTCCGGA	3842

RESULT 15

AAD37230
ID AAD37230 standard: DNA: 4182 BP.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

AC AAD37230;
XXDT 21-AUG-2002 (first entry)
yy

Human dystrophin minigene delta4

XX Human; dystrophin minigene; muscle

adeno-associated virus; AAV; Duck

KW Becker muscular dystrophy; ds.
XX

OS Homo sapiens.
XX

PN WO200183695-A2.

XX PD 08-NOV-2001.

[illegible]

27-APR-2001; 2001WU-US013677.
XX

PR 28-APR-2000; 2000US-0200777P.
XX

PA (XIAO/) XIAO X.

XX
PI Xiao X:

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	

DR WPL; 2002-049342/06.
XX

PT New dystrophin minigene for treating dystrophy comprises an N-terminal

PT dyscophy comprises all N-terminal rod repeats, H1 and H4 domains and

PT gene.
XX

PS Example 1; Page 43-44; 71pp; Eng

XX CC The present invention relates to

CC a dystrophin minigene. The minigene

CC terminal domains, rod repeats,
CC domains and cysteine-rich domain

invention also relates to a record comprising dystrophin minigene c

element. The dystrophin minigenes

CC control element, in a recombinant
CC useful for treating Duchenne mus-

CC dystrophy (BMD) in a mammalian s
CC dystrophin minigene delta4173 c

CC hinge H1 and rods R1, R2 and R3)

CC hinge H4 and CR domain) and 1104
XX

SQ Sequence 4182 BP; 1309 A; 927 C;

		Query March	62.2%; Score 1245; DB 6; Length 4182;
		Best Local Similarity	73.1%; Pred. No. 0;
		Matches 2001; Conservative	0; Mismatches 0; Indels 736; Gaps 2;
QY	1	GGCAGTTTCATTGATGTGAGAGTGAAGTAACCTCGGACCGTTATCAAAACAGCTTTAGAAGAA	60
DB	1000	GGCAGTTTCATTGATGTGAGAGTGAAGTAACCTCGGACCGTTATCAAAACAGCTTTAGAAGAA	1059
QY	61	GTAATACTCGTGCCCTTTCTGCTGAGGACACATTCGAAGCACAGAGAGAGATTCTAAT	120
DB	1060	GTAATACTCGTGCCCTTTCTGCTGAGGACACATTTGCAGACACAAGSAGAGATTCTAAT	1119
QY	121	GATGTGAAGTGTGTGAAGACCAGTTTCATCTCATGAGGGGTACATGATGGATTGACA	180
DB	1120	GATGTGAAGTGTGTGAAGACCAGTTTCATCTCATGAGGGGTACATGATGGATTGACA	1179
QY	181	GCCCATCAGGCCGCGTGTGGTATAATCTTCAAAATTTGGGAAGTGAAGCTGATTTGSAACAGGA	240
DB	1180	GCCCATCAGGCCGCGGTGGTATAATCTTCAAAATTTGGGAAGTGAAGCTGATTTGSAACAGGA	1239
QY	241	AAATTATTCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAAATCCAAGA	300
DB	1240	AAATTATTCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAAATCCAAGA	1299
QY	301	TGGGAATGCCCTCAGGCTAGCTAGCTAGTGAAGAAACAAGAACAAATTTACATAGAGTTTAAATG	360
DB	1300	TGGGAATGCCCTCAGGCTAGCTAGCTAGTGAAGAAACAAGAACAAATTTACATAGAGTTTAAATG	1359
QY	361	GATCTCCAGAATC-GAAACTGAAAGAGTTTGAATGACTGGCTAAACAAAAACAGAGAAAGA	419
DB	1360	GATCTCCAGAATCAGAAACTGAAAGAGTTTGAATGACTGGCTAAACAAAAACAGAGAAAGA	1419
QY	420	ACAAGGAATGTGGAGAGAGGCTCTTGACCTGATCTTTGAAGACCTAAACACGCCAAGTA	479
DB	1420	ACAAGGAATGTGGAGAGAGGCTCTTGACCTGATCTTTGAAGACCTAAACACGCCAAGTA	1479
QY	480	CACAAACATAGTGTCTTCAAGAAGATCTAGAACAAAGACAAGTCAAGGTCAAATCTCTCTC	539
DB	1480	CACAAACATAGTGTCTTCAAGAAGATCTAGAACAAAGACAAGTCAAGGTCAAATCTCTCTC	1539
QY	540	ACTCACATGTTGTTGTGTAGTAGTGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA	599
DB	1540	ACTCACATGTTGTTGTGTAGTAGTGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA	1599
QY	600	GAACAACCTTAAGTATTGGAGATCGATGGGCAAAATCTGTAGATGGACAGAGACCGC	659
DB	1600	GAACAACCTTAAGTATTGGAGATCGATGGGCAAAATCTGTAGATGGACAGAGACCGC	1659
QY	660	TGGGTTCTTTTACAGACATCTTTCTCAAATGGCAACGCTTTACTTGAAGAACAGTGCCTT	719
DB	1660	TGGGTTCTTTTACAGACATCTTTCTCAAATGGCAACGCTTTACTTGAAGAACAGTGCCTT	1719
QY	720	TTTTAGTCATGGCTTTTCAGAAAAAGAGATCAGTGAAACAAGATTCACACAACCTGGCTTT	779
DB	1720	TTTTAGTCATGGCTTTTCAGAAAAAGAGATCAGTGAAACAAGATTCACACAACCTGGCTTT	1779
QY	780	AAGATCAAATGAATGTTATCAAGTCTTCAAACCTGGCGTTTAAAAACGGGATCTA	839
DB	1780	AAGATCAAATGAATGTTATCAAGTCTTCAAACCTGGCGTTTAAAAACGGGATCTA	1839
QY	840	GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTTTTCACAA	899
DB	1840	GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTTTTCACAA	1899
QY	900	CTGAGAGATAGTCAGTGACCCAGAGACGGAGAGATGCTCGATTAACCTTCCC CGGTGT	959
DB	1900	CTGAGAGATAGTCAGTGACCCAGAGACGGAGAGATGCTCGATTAACCTTCCC CGGTGT	1959
QY	960	TGGGTAATTTAGTCCCCAAAACTTGAAAAAGAGTACAGCAC-----	999
DB	1960	TGGGTAATTTAGTCCCCAAAACTTGAAAAAGAGTACAGCACACTCATAGATTACTGC	2019

QY 1000 ----- 999
Db 2020 CAGTTCCTGGACCTGGAAAGTTTCTTGCTGGCTTACAGAGCTGAACAACACTGCC 2079
QY 1000 ----- 999
Db 2080 AATGCTCTACAGGATGTAACCGTAAGAAAGGCTCCTAGAGACTCCAAAGGAGTAA 2139
QY 1000 ----- 999
Db 2140 GAGCTGATGAACAATGGCAAGACCTCCAAGTGAATTTGAAGCTCACACAGATGTTAT 2199
QY 1000 ----- 999
Db 2200 CACAACCTGATGAACAACAGCCAAATAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 2259
QY 1000 ----- 999
Db 2260 GTCTGTTACAAGACGTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAG 2319
QY 1000 ----- 999
Db 2320 TCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCACTGGAAGGCTGTGCACCTT 2379
QY 1000 ----- 999
Db 2380 TCTCTGAGGAACCTTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCA 2439
QY 1000 ----- 999
Db 2440 CCTATTGAGCGACTTTCAGCAGTTTCAGAGCAGAACGATGATACATAGGGCCTTCAAG 2499
QY 1000 ----- 999
Db 2500 AGGAATTGAACCTAAGAAACCTGTAACTCATGAGTACTCTTGAGACTGTACGAATATT 2559
QY 1000 ----- 999
Db 2560 CTGACAGCAGCCTTTGGAGGACTAGAGAACTCTACAGAGGCCAGAGGCTGCCT 2619
QY 1000 ----- 999
Db 2620 CCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAAT 2679
QY 1000 -----AGACC 1004
Db 2680 ACTGAGTGGGAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 2739
QY 1005 CTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGATGAGCTGAGCTCAAGCTGGCCAA 1064
Db 2740 CTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGATGAGCTGAGCTCAAGCTGGCCAA 2799
QY 1065 GCTGAGGTGATCAAGGATCTGCGAGCCGCTGGGATCTCTCTATTGACTCTCTCCAA 1124
Db 2800 GCTGAGGTGATCAAGGATCTGCGAGCCGCTGGGATCTCTCTATTGACTCTCTCCAA 2859
QY 1125 GATCACCTCGAGAAAGTCAAGGACCTTCAGAGGAAATTCGCGCTCTGAAAGAGAACGTG 1184
Db 2860 GATCACCTCGAGAAAGTCAAGGACCTTCAGAGGAAATTCGCGCTCTGAAAGAGAACGTG 2919
QY 1185 AGCAGCTCAATGACCTTGTCTGCGAGCTTACCACTTTGGGATTCAGCTCTCACCGTAT 1244
Db 2920 AGCAGCTCAATGACCTTGTCTGCGAGCTTACCACTTTGGGATTCAGCTCTCACCGTAT 2979
QY 1245 AACTCAGCACTCTGGAAGACCTTGAACACAGATGAAGCTTCTGAGGTGGCTCGAG 1304
Db 2980 AACTCAGCACTCTGGAAGACCTTGAACACAGATGAAGCTTCTGAGGTGGCTCGAG 3039
QY 1305 GACCGAGTCAAGGAGCTGCAATGAAGCCCAAGGAACTTTGGTCCAGCACTCTCAGCACTTT 1364
Db 3040 GACCGAGTCAAGGAGCTGCAATGAAGCCCAAGGAACTTTGGTCCAGCACTCTCAGCACTTT 3099
QY 1365 CTTTCCAGCTCTGTCCAGGCTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCTTAC 1424

Db 3100 CTTTCCAGCTCTGTCCAGGCTCCCTGGGAGAGCCATCTCGCCAAACAAGTGCCTTAC 3159
QY 1425 TATATCAACACGAGACTCAAAACAACTTCTGGGACCATCCCAAAATGACAGAGCTTAC 1484
Db 3160 TATATCAACACGAGACTCAAAACAACTTCTGGGACCATCCCAAAATGACAGAGCTTAC 3219
QY 1485 CAGCTTTTAGTGTAGCTGAATAATGTCTCAGATTCTCAGCTTATAGGACTGGCATGAAATC 1544
Db 3220 CAGCTTTTAGTGTAGCTGAATAATGTCTCAGATTCTCAGCTTATAGGACTGGCATGAAATC 3279
QY 1545 CGAAGACTGCAAGAGGCCCTTTGCTTGGATCTCTTGGCCCTGTGAGCTGATGATGCC 1604
Db 3280 CGAAGACTGCAAGAGGCCCTTTGCTTGGATCTCTTGGCCCTGTGAGCTGATGATGCC 3339
QY 1605 TTGAGACGACAAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCGAGATTATTAAT 1664
Db 3340 TTGAGACGACAAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCGAGATTATTAAT 3399
QY 1665 TGTGACCACTATTTATGACCCCTGGAGCAAGACCAACAATTTGGTCAAGCTCCCT 1724
Db 3400 TGTGACCACTATTTATGACCCCTGGAGCAAGACCAACAATTTGGTCAAGCTCCCT 3459
QY 1725 CTCTGCGTGGATATGTCTGAACCTGCTGATGTTTATGATACGGGACGAACAGGG 1784
Db 3460 CTCTGCGTGGATATGTCTGAACCTGCTGATGTTTATGATACGGGACGAACAGGG 3519
QY 1785 AGGATCCGCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACAATTTGGAA 1844
Db 3520 AGGATCCGCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACAATTTGGAA 3579
QY 1845 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGGC 1904
Db 3580 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGGC 3639
QY 1905 AGGCTGGGCTCTCTTCTGCAATGATTTATCCAAATTCGAAGAGAGTGGGTGAAGTTGCA 1964
Db 3640 AGGCTGGGCTCTCTTCTGCAATGATTTATCCAAATTCGAAGAGAGTGGGTGAAGTTGCA 3699
QY 1965 TCCTTTGGGGCAGTAACATTTAGCCCAAGTGTCCGGA 2001
Db 3700 TCCTTTGGGGCAGTAACATTTAGCCCAAGTGTCCGGA 3736

Search completed: September 19, 2004, 04:49:36
Job time : 582.946 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 19, 2004, 03:01:40 ; Search time 103.19 Seconds
(without alignments)
10761.315 Million cell updates/sec

Title: US-09-845-416-14_COPY_1000_3000
Perfect score: 2001
Sequence: 1 gscagttcattgatgagag.....cattgagccaagtgtccgga 2001

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 27747546 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
5: /cgn2_6/prodata/2/ina/PTUS COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1002.6	50.1	5952	4	US-09-687-875A-1
2	921.6	49.6	13977	4	US-09-484-970B-60
3	869.8	43.5	19307	3	US-08-836-022A-10
4	869.8	43.5	19307	3	US-08-427-048A-10
5	490.6	24.5	6045	4	US-09-091-501B-7
6	490.6	24.5	10320	4	US-09-091-501B-9
7	460.6	23.0	3915	4	US-09-976-594-93
8	79.4	4.0	200	4	US-09-091-501B-5
9	78.6	3.9	200	4	US-09-091-501B-4
10	78.6	3.9	200	4	US-09-091-501B-6
11	74.6	3.7	7218	1	US-08-232-463-14
12	44.2	2.2	2574	4	US-09-668-313A-10
13	42.8	2.1	1179	4	US-09-107-532A-1186
14	42.8	2.1	1690	4	US-09-620-312D-69
15	42.8	2.1	7812	3	US-09-368-590-1
16	42.4	2.1	2704	3	US-08-857-076-44
17	42.4	2.1	3499	3	US-08-857-076-43
18	40.6	2.0	505	4	US-09-621-976-15639
19	39.4	2.0	832	4	US-09-621-976-2813
20	39.4	2.0	2223	1	US-08-257-073-4
21	39.4	2.0	193303	4	US-09-497-855A-37
22	39.4	2.0	193303	4	US-09-497-855A-44
23	38.8	1.9	289	3	US-09-007-005-17
24	38.8	1.9	289	3	US-09-244-796-17
25	38.8	1.9	1821	4	US-08-477-831C-1
26	38.8	1.9	1885	4	US-08-477-831C-9
27	38.8	1.9	1896	4	US-08-477-831C-10

28	38.8	1.9	1961	4	US-08-477-831C-8	Sequence 8, Appli
29	38.8	1.9	2968	4	US-08-477-831C-13	Sequence 13, Appl
30	38.8	1.9	3044	4	US-08-477-831C-12	Sequence 12, Appl
31	38.4	1.9	7672	4	US-09-220-132-24	Sequence 24, Appl
32	38.2	1.9	428	4	US-09-668-313A-3	Sequence 3, Appli
33	38.2	1.9	1848	4	US-09-134-001C-447	Sequence 447, App
34	38.2	1.9	4439	4	US-09-668-313A-17	Sequence 17, Appl
35	37.4	1.9	2082	3	US-08-985-335-4	Sequence 4, Appli
36	37.4	1.9	2082	3	US-08-410-372-4	Sequence 4, Appli
37	37.2	1.9	2160	4	US-09-052-218-1	Sequence 1, Appli
38	37.2	1.9	2915	4	US-09-336-115C-5	Sequence 5, Appli
39	37.2	1.9	3902	4	US-08-961-527-212	Sequence 212, App
40	36.8	1.8	4929	4	US-09-620-312D-674	Sequence 674, App
41	36	1.8	608	3	US-09-385-982-236	Sequence 236, App
42	36	1.8	2763	1	US-08-248-466B-2	Sequence 2, Appli
43	35.8	1.8	633	4	US-08-134-001C-578	Sequence 578, App
44	35.6	1.8	2447	2	US-09-014-869-14	Sequence 14, Appl
45	35.6	1.8	4868	1	US-08-139-937-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1

US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Paul
; APPLICANT: Xiao, Xiao
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPICED P
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687.875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/159,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match 50.1%; Score 1002.6; DB 4; Length 5952;

Best Local Similarity 99.6%; Pred. No. 1.4e-301;

Matches 1005; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	993	ACAGCACAGACCCCTGAAGACTCCAGAACTCCAGAGGCCACGGATGAGCTGACCTC 1052
DB	3679	ATAGATGAGACCCCTTGAAGACTCCAGAACTTCAGAGGCCACGGATGAGCTGACCTC 3738
QY	1053	AAGCTGCCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGCATCTCCTCAT 1112
DB	3739	AAGCTGCCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGCGCATCTCCTCAT 3798
QY	1113	GACTCTCCAGATCACCTCGAGAACTTCAGGCACTTCAGGAGAAATTCGCCTCTG 1172
DB	3799	GACTCTCCAGATCACCTCGAGAACTTCAGGCACTTCAGGAGAAATTCGCCTCTG 3858
QY	1173	AAAGAGAACGTGAGCCACGTCAATGACCTTCTGCCAGCTTACACATTTGGGCAATTCAG 1232
DB	3859	AAAGAGAACGTGAGCCACGTCAATGACCTTCTGCCAGCTTACACATTTGGGCAATTCAG 3918
QY	1233	CTCTCACCGTATAACCTCAGCACTTGGAACCTGGAACCTGGAACCTGGAACCTTCTGCAG 1292

Db 3919 CTCTCACCGTATACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAG 3978
Qy 1293 GTGGCCCTCGAGGACCGAGTCAGGACGCTCATGAGAGCCACACAGGACATTTGGTCCAGCA 1352
Db 3979 GTGGCCCTCGAGGACCGAGTCAGGACGCTCATGAGAGCCACACAGGACATTTGGTCCAGCA 4038
Qy 1353 TCTCAGCACTTTCTTTCCACGCTGTGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAAC 1412
Db 4039 TCTCAGCACTTTCTTTCCACGCTGTGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAAC 4098
Qy 1413 AAGTGCCTTACTATATCAACACGAGACTCAACAACTGTCTGGAGACATCCCAAAATG 1472
Db 4099 AAGTGCCTTACTATATCAACACGAGACTCAACAACTGTCTGGAGACATCCCAAAATG 4158
Qy 1473 ACAGAGCTTACACGCTTTTAGCTGACCTGAATAATGTCAAGATTTCTCAGCTTATAGACT 1532
Db 4159 ACAGAGCTTACACGCTTTTAGCTGACCTGAATAATGTCAAGATTTCTCAGCTTATAGACT 4218
Qy 1533 GCATGAACTCCGAAGACTGCAAGAGCCCTTTGTCTTGGATCTCTTTGAGCCCTGTGAGCT 1592
Db 4219 GCATGAACTCCGAAGACTGCAAGAGCCCTTTGTCTTGGATCTCTTTGAGCCCTGTGAGCT 4278
Qy 1593 GCATGTGATGCTTGTGAGGACGACACCTCAAGCAAAATGACACGACCCATGGATATCCTG 1652
Db 4279 GCATGTGATGCTTGTGAGGACGACACCTCAAGCAAAATGACACGACCCATGGATATCCTG 4338
Qy 1653 CAGATTATTAATTTGACCACTATTATGACCGCTGGAGCAAGAGACACAACTTTG 1712
Db 4339 CAGATTATTAATTTGACCACTATTATGACCGCTGGAGCAAGAGACACAACTTTG 4398
Qy 1713 GTCAAGCTTCTCTGCTGGATATGTCTGAACTGGCTGCTGAATGTTATGATAG 1772
Db 4399 GTCAAGCTTCTCTGCTGGATATGTCTGAACTGGCTGCTGAATGTTATGATAG 4458
Qy 1773 GGACGAACAGGAGGATCCCTGCTCTTTTAAACTGGCATCATTTTCCCTGTGTAA 1832
Db 4459 GGACGAACAGGAGGATCCCTGCTCTTTTAAACTGGCATCATTTTCCCTGTGTAA 4518
Qy 1833 GCACATTTGGAAGACAGTACAGTACCTTTTCAAGCAAGTGCGAAGTTCAACAGGATTT 1892
Db 4519 GCACATTTGGAAGACAGTACAGTACCTTTTCAAGCAAGTGCGAAGTTCAACAGGATTT 4578
Qy 1893 TGTGACGAGGAGGCTGGGCTCTCTCTGATGATTTTCAAAATTTCAAGACAGTTG 1952
Db 4579 TGTGACGAGGAGGCTGGGCTCTCTCTGATGATTTTCAAAATTTCAAGACAGTTG 4638
Qy 1953 GGTGAAGTTGCATCTTTTGGGGCAGTAAACATTTAGCCCAAGTGTCCGGA 2001
Db 4639 GGTGAAGTTGCATCTTTTGGGGCAGTAAACATTTAGCCCAAGTGTCCGGA 4687

RESULT 2

US-09-484-970B-60
; Sequence 60, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484, 970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 60
; LENGTH: 13977
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CBI
; NAME/KEY: unsure

; LOCATION: 11721-11761, 12294, 13969
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-60

Query Match 49.6%; Score 991.6; DB 4; Length 13977;
Best Local Similarity 99.5%; Pred. No. 6.8e-298;
Matches 1005; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 993 ACAGCAGACCCCTTGAAGACTCCAGAACTTCAAGAGCCACGAGATGAGCTGGACCTC 1052
Db 8993 ATAGATGAGACCCCTTGAAGACTCCAGAACTTCAAGAGCCACGAGATGAGCTGGACCTC 9052
Qy 1053 AAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCTTGGCAGCCCGTGGCGCATCTCTCAAT 1112
Db 9053 AAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCTTGGCAGCCCGTGGCGCATCTCTCAAT 9112
Qy 1113 GACTCTCTCAGATCACTTCAAGAACTCAAGGACCTTCAAGGAGAAATTTGGCCTCTG 1172
Db 9113 GACTCTCTCAGATCACTTCAAGAACTCAAGGACCTTCAAGGAGAAATTTGGCCTCTG 9172
Qy 1173 AAAGAGAACTGAGCCACGCTCAATGACTTGTCTGCCAGCTTACCACTTTGGGCAATTCAG 1232
Db 9173 AAAGAGAACTGAGCCACGCTCAATGACTTGTCTGCCAGCTTACCACTTTGGGCAATTCAG 9232
Qy 1233 CTCTCAGCTATTAACCTCAGACTCTGGAAGACTTGAACACCAAGATGGAAGCTTTGTCAG 1292
Db 9233 CTCTCAGCTATTAACCTCAGACTCTGGAAGACTTGAACACCAAGATGGAAGCTTTGTCAG 9292
Qy 1293 GTGGCCCTCGAGGACCGAGTCAGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCA 1352
Db 9293 GTGGCCCTCGAGGACCGAGTCAGCAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCA 9352
Qy 1353 TCTCAGCACTTTCTTTCCAGCTGTCTCCAGGTCCTTGGAGAGAGCCATCTCGCCAAAC 1412
Db 9353 TCTCAGCACTTTCTTTCCAGCTGTCTCCAGGTCCTTGGAGAGAGCCATCTCGCCAAAC 9412
Qy 1413 AAAGTGCCTTACTATATCAACACGAGACTCAACAACTTGTCTGGAGACCAATCCCAAAATG 1472
Db 9413 AAAGTGCCTTACTATATCAACACGAGACTCAACAACTTGTCTGGAGACCAATCCCAAAATG 9472
Qy 1473 ACAGAGCTTACCAAGCTTTTAGCTGACCTGAATAATGTCAAGATTTCTCAGCTTATAGACT 1532
Db 9473 ACAGAGCTTACCAAGCTTTTAGCTGACCTGAATAATGTCAAGATTTCTCAGCTTATAGACT 9532
Qy 1533 GCATGAACTCCGAAGACTGCAAGAGCCCTTTGTCTGATCTCTTGAAGCTTGTAGCT 1592
Db 9533 GCATGAACTCCGAAGACTGCAAGAGCCCTTTGTCTGATCTCTTGAAGCTTGTAGCT 9592
Qy 1593 GCATGTGATGCTTGGACACGACCACTCAAGCAAAATGACCAAGCCATGGATATCCTG 1652
Db 9593 GCATGTGATGCTTGGACACGACCACTCAAGCAAAATGACCAAGCCATGGATATCCTG 9652
Qy 1653 CAGATTATTAATTTGTTGACACTTATTTAGCCGCTGGAGCAAGAGCAACAATTTG 1712
Db 9653 CAGATTATTAATTTGTTGACACTTATTTAGCCGCTGGAGCAAGAGCAACAATTTG 9712
Qy 1713 GTCAAGCTTCTCTGCTGGATATGTCTGAACTGGCTGTGATGTTTATGATAG 1772
Db 9713 GTCAAGCTTCTCTGCTGGATATGTCTGAACTGGCTGTGATGTTTATGATAG 9772
Qy 1773 GGACGAACAGGAGGATCCCTGTCTTTTAAACTGGCATCATTTTCCCTGTGTAA 1832
Db 9773 GGACGAACAGGAGGATCCCTGTCTTTTAAACTGGCATCATTTTCCCTGTGTAA 9832
Qy 1833 GCACATTTGGAAGACAGTACAGTACCTTTTCAAGCAAGTGCGAAGTTCAACAGGATTT 1892
Db 9833 GCACATTTGGAAGACAGTACAGTACCTTTTCAAGCAAGTGCGAAGTTCAACAGGATTT 9892
Qy 1893 TGTGACGAGGAGGCTGGGCTCTCTCTGATGATTTTATCCAAATTTCAAGACAGTT 1951
Db 9893 TGTGACGAGGAGGCTGGGCTCTCTCTGATGATTTTATCCAAATTTCAAGACAGTT 9952
Qy 1952 GGTGAAGTTGCATCTTTTGGGGCAGTAAACATTTAGCCCAAGTGTCCGGA 2001

Db 9953 GGGTGAAGTTGATCTTGGGGGAGTAACTTGGAGCAAGTGTCTCGGA 10002

Db 5461 CTCTCACCTTATAACCTCAGCAGCTTTGGAAGATCTGAATACCAGATGGAGGCTTCTACAG 5402

Qy 1293 GTGGCCCTCGAGNACCGAGTCAAGCAGCTGATGAGCCGCCACAGGAGCTTTGGTCCAGCA 1352

Db 5401 GTGGCTGTGGAGGACCGTGTGACAGCTGATGAGCCGCCACAGGAGCTTTGGTCTTGA 5342

Qy 1353 TCTCAGCAGCTTTCTTCCAGCTGTCTCCAGGCTCCCTGGGAGAGAGCCATCTCCGCAAAAC 1412

Db 5341 TCCAGCAGCTTCTTCCAGCTGTCTCCAGGCTCCCTGGGAGAGAGCCATCTCCGCAAAAC 5282

Qy 1413 AAAGTGCCTTACTATATCAACAGCAGCTCAACAACTTCTGGGAGCCATCCCAAAATG 1472

Db 5281 AAAGTGCCTTACTATATCAACAGCAGCTCAACAACTTCTGGGAGCCATCCCAAAATG 5222

Qy 1473 ACAGAGCTTACAGCTCTTTAGCTGACCTGAATAATGTCAAGTCTCAGATCTCAGCTTATAGACT 1532

Db 5221 ACAGAGCTTACAGCTCTTTAGCTGACCTGAATAATGTCAAGTCTCAGATCTCAGCTTATAGACT 5162

Qy 1533 GCCATGAAGTCCGAGAGCTGAGAGCCCTTTGCTTGGATCTCTTTGAGCCTGTAGCT 1592

Db 5161 GCCATGAAGTCCGAGAGCTTCCAGAGGCTTCTTGGTGGATCTCTTTGAGCCTGTAGCT 5102

Qy 1593 GCATGTGATGCTTGGAGCAGCACAACCTCAAGCAAAATGACAGCCCATGGATATCTG 1652

Db 5101 GCATGTGATGCTTGGAGCAGCACAACCTCAAGCAAAATGACAGCCCATGGATATCTG 5042

Qy 1653 CAGATTATTAATGTTTGAACACTTATTTATGACCGCTGGAGAGAGCAACAATTTG 1712

Db 5041 CAGATAATTAATGTTTGAACACTTATTTATGACCGCTGGAGAGAGCAACAATTTG 4982

Qy 1713 GTCAAGCTCCCTCTCTGCTGGATATGTCTGAGCTGGCTGCTGATGTTTATGATACG 1772

Db 4981 GTCAAGCTCCCTCTCTGCTGGATATGTCTGAGCTGGCTGCTGATGTTTATGATACG 4922

Qy 1773 GGAGCAACAGGAGGATCCCTGTCTCTTTTAAACTGGCATCATTTCCCTGTGAAA 1832

Db 4921 GGAGCAACAGGAGGATCCCTGTCTCTTTTAAACTGGCATCATTTCCCTGTGAAA 4862

Qy 1833 GCATTTTGGAGAGAGTACAGATACCTTTTCAAGCAAGTGGAGTCAACAGGATTT 1892

Db 4861 GCACACTTGGAGAGAGTACAGATACCTTTTCAAGCAAGTGGAGTCAACAGGATTT 4802

Qy 1893 TGTGACGAGCGCAGGCTGGGCTCTCTCTGATGATTTCTTCAAAATTCAGAGAGTGTG 1952

Db 4801 TGTGACGAGCGTAGGCTGGGCTCTCTCTGATGATTTCTTCAAAATTCAGAGAGTGTG 4742

Qy 1953 GGTGAAGTGCATCTTTTGGGGGAGTAACTTGGAGCAAGTGTGCTCGGA 2001

Db 4741 GGTGAAGTGTCTTCTTTGGGGGAGTAACTTGGAGCAAGTGTGCTCGGA 4693

RESULT 4

US-09-427-048A-10/c

; Sequence 10, Application US/09427048A

; Patent No. 6203975

; GENERAL INFORMATION:

; APPLICANT: Trustees of the University of Pennsylvania

; Wilson, James M.

; Fisher, Krishna J.

; Chen, Shu-Jen

; Weitzman, Matthew

; TITLE OF INVENTION: Improved Adenovirus Virus and Methods of Use Thereof

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Howson and Howson

; STREET: Spring House Corporate Cntr, P O Box 457

; CITY: Spring House

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19477

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

Db 9953 GGGTGAAGTTGATCTTGGGGGAGTAACTTGGAGCAAGTGTCTCGGA 10002

Db 993 ACAGCAGACCTTGAAGACTCCAGGACTTCAAGAGCCAGGATGAGTGGACCTC 1052

Db 5701 ATAGATGAAGCTTTGAAGACTCCAGGACTTCAAGAGCCAGGATGAGTGGACCTC 5642

Qy 1053 AAGTGTGCGCAAGCTGAGTGTATCAAGGATCTTGGAGCCCTGGGCGATCTCTCAT 1112

Db 5641 AAGTGTGCGCAAGCTGAGTGTATCAAGGATCTTGGAGCCAGTGGGCGATCTCTCAT 5582

Qy 1113 GACTCTTCAAGATCACTCGAGAGTCAAGGACTTGGAGCCCTGGGCGATCTTGGGCGATCT 1172

Db 5581 GACTCTTCAAGATCACTCGAGAGTCAAGGACTTGGAGCCCTGGGCGATCTTGGGCGATCT 5522

Qy 1173 AAAGAGAACTGAGCAGCTCAATGACTTGTCTGGAGCTTACCACTTTGGGCGATCT 1232

Db 5521 AAAGAGAACTGAGCAGCTCAATGACTTGTCTGGAGCTTACCACTTTGGGCGATCT 5462

Qy 1233 CTCTCAGCTATACCTTCAAGCTTGGAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAG 1292

Query Match 43.5%; Score 869.8; DB 3; Length 19307;

Best Local Similarity 91.4%; Pred. No. 8.3e-260;

Matches 922; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy 993 ACAGCAGACCTTGAAGACTCCAGGACTTCAAGAGCCAGGATGAGTGGACCTC 1052

Db 5701 ATAGATGAAGCTTTGAAGACTCCAGGACTTCAAGAGCCAGGATGAGTGGACCTC 5642

Qy 1053 AAGTGTGCGCAAGCTGAGTGTATCAAGGATCTTGGAGCCCTGGGCGATCTCTCAT 1112

Db 5641 AAGTGTGCGCAAGCTGAGTGTATCAAGGATCTTGGAGCCAGTGGGCGATCTCTCAT 5582

Qy 1113 GACTCTTCAAGATCACTCGAGAGTCAAGGACTTGGAGCCCTGGGCGATCTTGGGCGATCT 1172

Db 5581 GACTCTTCAAGATCACTCGAGAGTCAAGGACTTGGAGCCCTGGGCGATCTTGGGCGATCT 5522

Qy 1173 AAAGAGAACTGAGCAGCTCAATGACTTGTCTGGAGCTTACCACTTTGGGCGATCT 1232

Db 5521 AAAGAGAACTGAGCAGCTCAATGACTTGTCTGGAGCTTACCACTTTGGGCGATCT 5462

Qy 1233 CTCTCAGCTATACCTTCAAGCTTGGAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAG 1292

; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/427,048A
 ; FILING DATE: 21-Oct-1999
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/836,022
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bak, Mary E.
 ; REGISTRATION NUMBER: 31,215
 ; REFERENCE/DOCKET NUMBER: GNVPM.008PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 215-540-9200
 ; TELEFAX: 215-540-5818
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 19307 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: cDNA
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
 US-09-427-048A-10

Query Match 43.5%; Score 869.8; DB 3; Length 19307;
 Best Local Similarity 91.4%; Pred. No. 8.3e-260;
 Matches 922; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY	993	ACAGCAGACCCCTTGAAGACTCCAGGAACTTCAAGAGCCACGGATGAGTGGACCTC	1052
DB	5701	ATAGATGAAGCTCTTGAAGACTCCAGGAACTTCAAGAGCTGCCATGAATGGACCTC	5642
QY	1053	AAGCTGGCCCAAGCTGAGGTGATCAAGGATCTCTGGCAGCCCTGGGCGATCTCTCAT	1112
DB	5641	AAGTGGCCCAAGCTGAGGTGATCAAGGATCTCTGGCAGCCCTGGGCGATCTCTCAT	5582
QY	1113	GACTCTTCCAGATCACCTCGAAGAACTCAAGGCATTCGAGGAAATTCGGCTCTG	1172
DB	5581	GACTCTTCCAGATCACCTTGAAGAACTCAAGGCATTCGAGGAAATTCGACCTCTT	5522
QY	1173	AAAGAGACGTGAGCCACGTCAATGACCTTGTCCGACCTTACCATTTGGGCATTG	1232
DB	5521	AAAGAGATGTCAATGTGTCATGACCTTGTCCGACCTTACCATTTGGGCATTG	5462
QY	1233	CTCTCACCGTATAACCTCAGCCTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGCAG	1292
DB	5461	CTCTCACCTTATAACCTCAGCCTTGGGAAGATCTGAATACCAGATGGAGGCTTCTACAG	5402
QY	1293	GTGGCGCTCAGGACCGAGTCAAGGAGCTGATGAAGCCACAGGACCTTTGGTCAGCA	1352
DB	5401	GTGGCTGTGGAGGACCGTGTCAAGAGCTGATGAAGCCACAGGACCTTTGGTCGCA	5342
QY	1353	TCTCAGCAGCTTTCTTTCCAGCTGTGTCAGGCTCTGGAGAGAGCCATCTCGCAAAAC	1412
DB	5341	TCCAGCAGCTTTCTTTCCAGCTGTGTCAGGCTCTGGAGAGAGCCATCTCAACAAAC	5282
QY	1413	AAAGTGCCTTACTATATCAACACAGAGACTCAAAACAACTTGTGGGACCATCCCAAAATG	1472
DB	5281	AAAGTGCCTTACTATATCAACACAGAGACTCAAAACAACTTGTGGGACCATCCCAAAATG	5222
QY	1473	ACAGAGCTCTACAGCTTTTAGTGTACCTGATATGTCAGATTTCTCAGCTTATAGGACT	1532
DB	5221	ACAGAGCTCTACAGCTTTTAGTGTACCTGATATGTCAGTTCCTCCGCTATAGGACT	5162
QY	1533	GCCATGAACTCCGAAGACTGCGAAGAGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCT	1592
DB	5161	GCCATGAACTCAGAAGGCTCCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCT	5102
QY	1593	GCATGTGATGCTTGGACACAGACCACTTCAAGCAAAATGACACGCCATCGATATCTG	1652

DB	5101	GCATGTGATGCTTGGACACAGCAACCTCAAGCAAAATGACCAGCCCATGGATATCTCTG	5042
QY	1653	CAGATTATTAATGTTTGGACCACTATTATGACCGCTTGGAGCAAGGACACAAATTTG	1712
DB	5041	CAGATAATTAATGTTTGGACCACTATTATGATCGCTGGAGCAAGGACACAAATCTG	4982
QY	1713	GTCAACGTCCTCTCTGGTGGATATGCTGAACCTGCTGAATGCTTATGATACG	1772
DB	4981	GTCAATGCTCTCTCTGGTGGATATGCTGAACCTGCTGAATGCTTATGATACG	4922
QY	1773	GGAGCAACAGGAGGATCCGCTGCTCTTTTAAACTGGCATCATTTCCCTGTGFAAA	1832
DB	4921	GGAGCAACAGGAGGATCCGCTGCTCTTTTAAACTGGCATCATTTCCCTGTGFAAA	4862
QY	1833	GCACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATT	1892
DB	4861	GCACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATT	4802
QY	1893	TGTACCAAGCAGGCTGGGCTCTCTTCTGCATGATTTCTATCCAAATTCGAAGCAGTTG	1952
DB	4801	TGTACCAAGCAGGCTGGGCTCTCTTCTGCATGATTTCTATCCAAATTCGAAGCAGTTG	4742
QY	1953	GCTCAAGTTCATCTTTGGGGCAGTAAACATTTGAGCCCAAGTGTCCGGA	2001
DB	4741	GCTCAAGTTCATCTTTGGGGCAGTAAACATTTGAGCCCAAGTGTCCGGA	4693

RESULT 5
 US-09-091-501B-7
 ; Sequence 7, Application US/09091501B
 ; Patent No. 6518413
 ; GENERAL INFORMATION:
 ; APPLICANT: Tinsley, Jonathan M
 ; APPLICANT: Davies, Kay E
 ; TITLE OF INVENTION: Utrophin gene expression
 ; FILE REFERENCE: 620-42
 ; CURRENT APPLICATION NUMBER: US/09/091,501B
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: PCT/GB96/03156
 ; PRIOR FILING DATE: 1996-12-19
 ; PRIOR APPLICATION NUMBER: GB 9525962.8
 ; PRIOR FILING DATE: 1995-12-19
 ; PRIOR APPLICATION NUMBER: GB 9615797.9
 ; PRIOR FILING DATE: 1996-07-26
 ; PRIOR APPLICATION NUMBER: GB 9622174.2
 ; PRIOR FILING DATE: 1996-10-24
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 7
 ; LENGTH: 6045
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (11)..(6037)
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (724)..(758)
 ; OTHER INFORMATION: Precise residue is left open
 US-09-091-501B-7

Query Match 24.5%; Score 490.6; DB 4; Length 6045;
 Best Local Similarity 67.6%; Pred. No. 6.8e-142;
 Matches 688; Conservative 0; Mismatches 329; Indels 0; Gaps 0;

QY	985	AAAAGAGTACAGCAGACAGCCCTTGAAGAGCTCCAGGAACCTCAAGAGCCACCGATGAGC	1044
DB	3783	AAAAGCAAGTGGCAAGGCATTGGAGAACTCAGAGACCTTGCAGGAGCTATGATGACC	3842
QY	1045	TGACCTCAAGCTGCGCCCAAGCTGAGGTATCAAGGGATCTTGGCAGCCGCTGGCGATC	1104

Db 3843 TGGACGCTGACATGAGGAGGAGAGTCCGTCGGGATGGCTGGAAGCCCGTGGGAGACT 3902
Qy 1105 TCTCATTTGACTCTCTCCAGATCACTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTTG 1164
Db 3903 TACTCATTGACTCGCTGCGAGGATCAATTAAGAAATCATGCAATTTAGAGAAATTTG 3962
Qy 1165 CGCTCTGAAAGAGAGAGTGGAGCAGTCAATGACCTTCTCGCAGCTTACACTTTGG 1224
Db 3963 CACCAATCAACTTAAAGTTAAACCGTGATGATTTATCCAGTCAGCTGTCTCCACTTG 4022
Qy 1225 GCATTGAGTCTCACCGTATAAAGTCAAGCTCTGGAAGAGCTGGAACACAGATGAAGC 1284
Db 4023 ACCTGATCCCTCTCTAAAGATGCTCGCAGCTAGATGACCTTAATATGCGATGGAAC 4082
Qy 1285 TTCTGAGGTGGCGCTGAGGAGCGAGTCAAGGAGCTGATGAAGCCACAGGACTTTG 1344
Db 4083 TTTTACAGGTTTCTGTGGATGATCGGCTTAAACAGCTTCAGGAGCCACAGATTTG 4142
Qy 1345 GTCCAGATCTCAGCACTTTCTTCCAGCTCTCTCCAGGTCCTCGGAGAGGCAATCT 1404
Db 4143 GACCATCTCTCAGCACTTTCTCTACGTCAGTCCAGCTCGCTGGCAAGATCCATTT 4202
Qy 1405 CGCCAAACAAAGTCCCTACTATATCAACACAGAGACTCAACAACTTCTGAGCAATC 1464
Db 4203 CACATAATAAGTSCCTTATACATCAACCATCAAAACAGACCACTGTGGGACCATC 4262
Qy 1465 CCAAAATGACAGAGCTCTACCACTCTTTAGCTGACCTGAATATGTGAGATTTCTAGCT 1524
Db 4263 CTAAATGACGAGCTTTCTCAATCTCTGCTGACCTGATATGTACGTTTTCTGCTC 4322
Qy 1525 ATAGGACTGCCATGAAGTCCGAGAGCTGAGAGGCGCTTCTGCTGGATCTCTTGAGCC 1584
Db 4323 ACCGTACAGCAATCAAAATCCGAGAGCTACAAAGACACTATCTTTGGATCTCTTAGAT 4382
Qy 1585 TGTGAGTGCATGTGATGCTTGGACGAGCACTCAAGCAACCTCAAGCAAAATGACGCGCATGG 1644
Db 4383 TGAGTCAACAAATGAATTTTCAACAGCACAAGTTGAACCAAAATGACGCTCTCTCA 4442
Qy 1645 ATATCTGTCAGATTAATTAATTTGTTGACCACTATTATGACCGCTCGGAGCAAGACACA 1704
Db 4443 GTGTCCAGATGTCATCACTGCTGACAACTTATGATGAGCTTGACAAATGCATA 4502
Qy 1705 ACATTTGGTCAAGGCTCTCTGCTGATATGCTGCACTGCTGCACTGCTGAATTTT 1764
Db 4503 AGGACCTGGTCAAGCTTCCACTCTGTTGATATGTCTCAATTTGGTCTCAATGCT 4562
Qy 1765 ATGATCGGACGAGCAAGGAGGATCGTCTGCTCTTTTAAACTGGCATCATTTCCG 1824
Db 4563 ATGACAGGCTCGAACTTGAATTTAGATGAGAGCTGAGATTTGATTAATGCTC 4622
Qy 1825 TGTGTAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAA 1884
Db 4623 TCTCCAAAGGCTCTTTGGAAGAAAATACAGATATCTCTTTAAGGAAGTTGCGGGCCGA 4682
Qy 1885 CAGGATTTGTGACCGAGGCTGGGCTCTCTGCTGATGATTTCTATCCAAATTCOA 1944
Db 4683 CAGAAATGTGACAGAGGAGCTGGGCTGTTACTTCAATGATGCAATCCAGATCCCC 4742
Qy 1945 GACAGTTGGTGAAGTTGCATCTTTTGGGGGAGTAAATGAGCAAGTGTCCGGA 2001
Db 4743 GGCACTAGGTGAAGTAGCAGCTTTTGGAGGAGTAAATTTGAGCTAGTGTTCGCA 4799

RESULT 6

US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Ray E
; TITLE OF INVENTION: Trophic gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091.501B

; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 9
; LENGTH: 10320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(10312)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Full length
; OTHER INFORMATION: utrophin construct
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
; US-09-091-501B-9

Query Match 24.5%; Score 490.6; DB 4; Length 10320;
Best Local Similarity 67.6%; Pred. No. 9.9e-142;
Matches 588; Conservative 0; Mismatches 329; Indels 0; Gaps 0;
Qy 985 AAAAGAGTACAGCAGACACCTTGAAGACTTCCAGGAACTTCAAGAGCCACGATGAGC 1044
Db 8058 AAAAGCAAGTGGCAAGGCATTGGAGAACTCAGAGACCTCAGGAGACTATGATGACC 8117
Qy 1045 TGGACCTCAAGTCGCGCAAGCTGAGGTGATCAAGGGATCTGGCAGCCCTGGCGATC 1104
Db 8118 TGGACGTGACATGAGAGGACAGAGTCCGTCGGGATGGCTGGAGCCCTGGGAGACT 8177
Qy 1105 TCTCATTTGACTCTCTCCAGATCACTTCGAGAAAGTCAAGCACTTCGAGGAGAAATTTG 1164
Db 8178 TACTCATTGACTCGCTGCGAGGATCACATTTGAAATAATCATGGCACTTTAGAGAAATTTG 8237
Qy 1165 CGCTCTGAAAGAGAACTGAGCCACGTCATGACCTTGTCTGCGCAGCTTACCCTTTGG 1224
Db 8238 CACCAATCAACTTTAAAGTTAAACGGTGAATGATTTATCCAGTCAGCTGTCTCCACTTG 8297
Qy 1225 GCATTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAAGATGGAAGC 1284
Db 8298 ACTGCTATCCCTCTCTAAAGATGTCTCGCAGCTAGATGACCTTAATATGATGGAAC 8357
Qy 1285 TTCTCGAGTGGCCCTCGAGGACCGAGTCAGGAGCTGCTCATGAAGCCACAGGAGACTTTG 1344
Db 8358 TTTTACAGGTTTCTGTGGATGATCGCTTAAACAGCTTCAGGAAGCCACAGAGATTTG 8417
Qy 1345 GTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGTCAGGGTCCCTGGGAGAGAGCCATCT 1404
Db 8418 GACCATCTCTCAGCACTTTCTCTCTAGTCAGTCCAGCTGCGCTGGCAAGATCCATT 8477
Qy 1405 CGCCAAACAAAGTGCCTTACTATATCAACACAGAGACTCAAAACAACTTGTCTGGGAGCAATC 1464
Db 8478 CACATAATAAGTGCCTTATTACATCAACCATCAAAACAGACCAACCTGTGTGGGAGCAATC 8537
Qy 1465 CCAAAATGACAGAGCTCTACCACTTTTGTGCTGACCTGAATTAATGCTAGATTTCTCAGCTT 1524
Db 8538 CTAAATGACCGAATCTTTCAATCCCTTGTCTGACCTGAATTAATGATGCTTTTCTGCTC 8597
Qy 1525 ATAGGACTGCCATGAAACTCCGAGAGCTGCAAGAGCCCTTTGCTTGGATCTCTTGAAGCC 1584
Db 8598 ACCGTACAGCAATCAAAATCCGAGAGCTACAAAGAGCACTATGTTGGATCTCTTAGAGT 8657
Qy 1585 TGTGAGTGCATGTGATGCTTGGACGAGCAACCTCAAGCAAAATGACAGCCCATGG 1644

Db 8658 TGATACAAACAAATGAAATTTTCAACACAGCACAGTGTGAACAAATGACCAAGCTCCTCA 8717
Qy 1645 ATATCTGCAGATTAATTAATTTGACCACTATTATGACCGCTGGAGCAAGACACA 1704
Db 8718 GTGTTCAGATGTCATCAACTGTCGACAAACAATTTATGATGACATGAGCAATGCATA 8777
Qy 1705 ACAATTTGGTCAACGTCCTCTCTGCGTGATATGTCGAACTGGCTGCTCAATGTTT 1764
Db 8778 AGGACCTGGTCAAGTTCCTCTGTTGATATGTCCTCAATGGTGGCTCAATGTCT 8837
Qy 1765 ATGATAGCGACGACAAAGGAGGAGTCGTCCTCTCTGATGTTTAAACATGGCATCTTCC 1824
Db 8838 ATGACAGGGTCGAACTCGGAAAAATTAGATGCGAGATCTGAAGATTGGATTAAATGCTC 8897
Qy 1825 TGTGTAAGCACATTTGGAAGACAGATACAGATACCTTTTCAAGCAAGTGGCAAGTTCAA 1884
Db 8898 TCTCAAAGGTCTCTTGGAAAGAAAAATACAGATATCTCTTTAAGGAAGTTGGGGCCGA 8957
Qy 1885 CAGGATTTGTGACCGCAGCGCTGGCCCTCTCTGCGATGTTCTATCCAAATTCGAA 1944
Db 8958 CAGAAATGTGTGACCGCAGCGCTGGCCCTGTTTACTTCAATGATGCCATCCAGATCCCC 9017
Qy 1945 GACAGTTGGGTGAAGTTGCATCTTTGGGGGCGAGTAACATGAGCCAGTTCGGA 2001
Db 9018 GGCAGCTAGGTGAAGTAGCAGCTTTGGAGGCGAGTAATATTAGCCCTAGTGTTCGCA 9074

RESULT 7

US-09-976-594-93
; Sequence 93, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; CURRENT APPLICATION NUMBER: US/09/976,594
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 93
; LENGTH: 3915
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 290344.1
US-09-976-594-93

Query Match 23.0%; Score 460.6; DB 4; Length 3915;
Best Local Similarity 57.4%; Pred. No. 1.1e-132;
Matches 652; Conservative 0; Mismatches 319; Indels 0; Gaps 0;
Qy 1031 GGCCACGGATGAGTGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCA 1090
Db 1022 GGGATGAGAACTAAGCACTACTTGGCCCAAGCTGAGGTGATCGGAGTCCGAGCCACTGGGA 1081
Qy 1091 GCGCGTGGCGATCTCTCATTTGACTCTCTTCCAGATCCTCGAGAAAGTCAAGGCAT 1150
Db 1082 GCGCATGGGATCTCTTATTGATTCATCTCCAGAGCATCCAGGCTATTAAAGTGT 1141
Qy 1151 TCAGAGGAATTTGGCCTCTGAAGAGACGTCGAGCCAGTCATGACCTTCTCTCGCA 1210
Db 1142 CAAAGAGAAATTTCCCGCATGAAGATGGAGTAAAGTTGGTGAATGATCTGGCCCA 1201
Qy 1211 GCTTACCATTTGGGATTCAGCTCTCACCGTATAACCTCAGCACCTCTGGAAGACCTGAA 1270
Db 1202 ACTTGCCATTTCTGATGTGCATCTGTCAATGGAGATTTCCAGGCCCTGGAACAGATCAA 1261
Qy 1271 CACAGATGAAGCTCTGAGGTGGCGCTGAGGACCGAGTCAGCAGCTGCATGAAGC 1330

Db 1262 CGTCCGATGGAACAACTACAGGCGTCAGTTGATGAGAGGCTTAAGCAGCTCCAGGATGC 1321
Qy 1331 CCACAGGACATTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTCTCAGGGTCCCTG 1390
Db 1322 CCACGGGACTTTGGGCGCTGGGTCAAGCACTTTCTCTCTCTCTCTCTCCAGGTTCCCTG 1381
Qy 1391 GGAGAGACCATCTCGCCAAACAAAGTGCCTTACTATATCAACACAGAGCTCAAAAC 1450
Db 1382 GGAAAGACCAATTTCAACCAATTAAGTTCCTACTATCAATCAACACAGAGCTGAGCCAC 1441
Qy 1451 TTGCTGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAAATATGT 1510
Db 1442 ATGCTGGACCATCCCAAGATGACAGAGTTATACCAAACTTAGCTGATCTGAACACAT 1501
Qy 1511 CAGATTTCTCAGCTTATAGGCTGCATGAACTCCGAGAGCTGAGAGGCTCCAGAGCCCTTTCCTT 1570
Db 1502 TAAGTTCTCAGCTTATCGCTGCTGCAATGAACTCCGAGAGTCCAGAAAGCCCTGCGCTT 1561
Qy 1571 GGATCTCTTGAGCCTGTGAGCTGATGTGATGCTTGGACAGCACAACTCTCAGCAAAA 1630
Db 1562 GGACTGTGTAATTTAAACACAGCCCTGGAAATCTTCAATGAGCATGATCTGAGGCCAG 1621
Qy 1631 TGACAGCCCATGGAATCTCGAGATTAATTAATTTGACCACTATTTATGACCCGCT 1690
Db 1622 TGACAGCTGATGATGTGTGAGAGTCAATCACTGCTGACTGCTGCTTATATGAACGTTT 1681
Qy 1691 GGACAGAGCACAACTTTGGTCAACGCTCCCTCTCTGGTGATATGTCTGAACTG 1750
Db 1682 GGAGGAAAGAGGATCTGCTGCACTGCCACTCTGTGACATGAGCTCAATG 1741
Qy 1751 GCTGCTGAATTTTATGATCGGACGAAAGAGGAGTCCGTGCTCTCTTTTAAAC 1810
Db 1742 GCTGCTCAATGTTTGTGATGTCGCGGAAAGATGCGGCACTTGTCTTTAAGAC 1801
Qy 1811 TGGCATCATTTCCCTGTGTAAGACATTTGGAGACAACTGACATGATCTTTCAAGCA 1870
Db 1802 TGGATTTGATGCTTTGTGCGACGGAAGTGAAGAAACTTCACTGATCTTTCAGCCA 1861
Qy 1871 AGTGGCAAGTTCAACAGGATTTGTGACAGCGAGCTGGGCTCTCTGATGATTC 1930
Db 1862 AGTGGCAACTCAGGCGAGCTGTGACAGCGCACTTGTGCTCTCTCTTATGAGGC 1921
Qy 1931 TATCCAAATTCAGACAGTTGGTGAAGTGCATCTTTGGGGGCGAGTCAACTGAGCC 1990
Db 1922 CATTGAGTGGCGCTCAGCTGGGTGAAGTGGCAGCCCTTTGGGGGCGAGCAATGTGGAGCC 1981
Qy 1991 AAGTGTCCGGA 2001
Db 1982 CAGTGTCCGTA 1992

RESULT 8

US-09-091-501B-5
; Sequence 5, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5


```
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Rattus sp.
US-09-091-501B-5

Query Match      4.0%; Score 79.4; DB 4; Length 200;
Best Local Similarity 64.3%; Pred. No. 1.1e-14;
Matches 119; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 464 CTTAAACGCCAAGTACAACAACATAAAGTGTCTTCAAGAAGATCTAGAACAAACAAGT 523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16 CCTGCAAAACCTGCTTGAAGAACATAAAAGTTTGCAAGTGACCTCGAAGCTGACAGT 75

QY 524 CAGGTCAATCTCTCACTACATCATGTGTGGTAGTGTATGATCTAGTGGAGATCACGC 583
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 GAAGTGCAATCTCTTAACATCATATGTGTGTGATTGGATGAAACAGTGGGGAGAGCGC 135

QY 584 AACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATCATGGGCAACATCTGTAG 643
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 CACAGCTGTTTGGAGAGATCAGTTACAGAACTGGGTGAGCGCTGGACAGCTGTATGCCG 195

QY 644 ATGGA 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 CTGGA 200
```

```
RESULT 9
US-09-091-501B-4
; Sequence 4, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-091-501B-4

Query Match      3.9%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 1.9e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 452 TGATCTTGAAGACCTAAACGCCAAGTACAACAACATAAAGTGTCTTCAAGAAGATCTAGA 511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4 TGATGTGAATCTCTTCAAAAGCTGCTAGAGAACAATAAAGTTTGCAAGATGATCTTGA 63

QY 512 ACAAGAACAGTCAGGTCAATCTCTCACTACATCATGTGTGTGTAGTGAATCTAG 571
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 64 GGCTGAACAGGTGAAAGTAAATTCATACTACATCATGTGTGTGTATGATGAAACAG 123

QY 572 TGGAGATCACCAACTGCTGCTTTGGAGAACAACTTAAGTATTGGAGATCGATGGC 631
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 TGGTGAGAGCGCTACAGCTATCTTAGAAGACCACTTACAGAACTTGGTGAGCGCTGAC 183

QY 632 AAACATCTGTAGATGGA 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 184 AGCAGTATGCCGCTTGA 200
```

```
RESULT 10
US-09-091-501B-6
; Sequence 6, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-091-501B-6

Query Match      3.9%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 1.9e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 452 TGATCTTGAAGACCTAAACGCCAAGTACAACAACATAAAGTGTCTTCAAGAAGATCTAGA 511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4 TGATGTGAATCTCTTCAAAAGCTGCTAGAGAACAATAAAGTTTGCAAGATGATCTTGA 63

QY 512 ACAAGAACAGTCAGGTCAATCTCTCACTACATCATGTGTGTGTAGTGAATCTAG 571
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 64 GGCTGAACAGGTGAAAGTAAATTCATACTACATCATGTGTGTGTATGATGAAACAG 123

QY 572 TGGAGATCACCAACTGCTGCTTTGGAGAACAACTTAAGTATTGGAGATCGATGGC 631
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 TGGTGAGAGCGCTACAGCTATCTTAGAAGACCACTTACAGAACTTGGTGAGCGCTGAC 183

QY 632 AAACATCTGTAGATGGA 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 184 AGCAGTATGCCGCTTGA 200
```

```
RESULT 11
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
```

```
;
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
; US-08-232-463-14

Query Match          3.7%; Score 74.6; DB 1; Length 7218;
Best Local Similarity 5.6%; Pred. No. 4.3e-12;
Matches 23; Conservative 236; Mismatches 150; Indels 0; Gaps 0;

Qy 113 TTCTATATGATGTGGAAGTGGTGAAGACCCAGTTTCATATCTATGAGGGGTACATGATGG 172
Db 1474 TATCTATGCAAGTAGTTAAAGATAGACAGAAATTTGCTACRERRRRRRRRRRRRR 1415

Qy 173 ATTTCAGACCCATCAGGCGCGGGTGGTAATTTCTACAATTTGGAAGTAAGCTGATG 232
Db 1414 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1355

Qy 233 GAACAGGAAAATTATCAGAGATGAGAACTGAAGTACAAGACAGATGAATCTCTCAA 292
Db 1354 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1295

Qy 293 ATTCAAGATGGAAATGCTCAGGTAGCTAGCTAGTGAAGAAAACAAACAACTTACATAG 352
Db 1294 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1235

Qy 353 TTTTAATGATCTCCGAATCGAACTGAAGAGTTGAATGACTGCTAACAACACAGA 412
Db 1234 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1175

Qy 413 AGAAAGAACAGGAAATGAGGAGAGCGCTCTTGACCTGATCTTTGAAGACCTAAACG 472
Db 1174 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1115

Qy 473 CCAAGTACACACATAAGTGTCTTCAAGAGATCTAGACAGACAA 521
Db 1114 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1066

RESULT 12
US-09-668-313A-10
; Sequence 10, Application US/09668313A
; Patent No. 6503756
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRESSION
; FILE REFERENCE: RTS-0127
; CURRENT APPLICATION NUMBER: US/09/668,313A
; CURRENT FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 247
; SEQ ID NO 10
; LENGTH: 2574
```

```
;
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (218)...(1891)
; US-09-668-313A-10

Query Match          2.2%; Score 44.2; DB 4; Length 2574;
Best Local Similarity 54.7%; Pred. No. 0.0062;
Matches 88; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 1389 TGGGAGAGAGCCATCTGCCCAACAAAGTGCCCTACTATATCAACACAGAGACTCAAACA 1448
Db 1733 TGGGAGGAAGCTTACACAGCAGATGGAAATCAAGTACTTCAACACACGTCGACACAGACC 1792

Qy 1449 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTTACCAGTCTTTAGCTGACCTGAATAA 1508
Db 1793 ACGTCTGGATCCACCCCGTGATGAGCGCCCTGAACCTGCTCTGTGCGAGAGAGAGTGAA 1852

Qy 1509 GTCAGATCTCAGCTTATAGGACTGCCATGAAACTCCGAG 1549
Db 1853 GAGACTGTCCAGAGAGCTAACACAGCCCGAAAGCTGATG 1893

RESULT 13
US-09-107-532A-1186
; Sequence 1186, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: STC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1186:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1179 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
```

LOCATION: (B) LOCATION 1...1179
SEQUENCE DESCRIPTION: SEQ ID NO: 1186;
US-09-107-532A-1186

Query Match 2.1%; Score 42.8; DB 4; Length 1179;
Best Local Similarity 49.5%; Pred. No. 0.0097;
Matches 110; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 353 TTTTATGATCTCCAGATCGAACTGAAAGATTGATGCTGCTAACAACACAGA 412
DB 585 TGTGTATCCAGCAGTTGAGCAAGAATAAAGATTTCATGATCAGACAAAGAAATGG 644
QY 413 AGAAGAACAAGAAATGGAGAGAGCCTCTTTGACCTGATCTTTGAAGACCTTAAACG 472
DB 645 AGATACGATCGGAGGAATGTAGAACTGCTCGTTGAGCGCTTCCAGCTGGATTAGGAAG 704
QY 473 CCAAGTACACACATAGGTGCTTCAAGAGATCTAGAACAGACAAAGTCAGGGTCAA 532
DB 705 CTACGTACATGGGACACGAAGTAGATGCCAAATCGCACAAAGCTGTGTTAGTATCAA 764
QY 533 TTCTCTCACTCACATGCTGTGCTGATGATGAATCTAGTGG 574
DB 765 TGCCTTTAAAGCGGTAGAAATTTGGGTCGGATTCACCTTCTGG 806

RESULT 14
US-09-620-312D-69
; Sequence 69, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chonghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungling
; APPLICANT: Wang, Duanui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radjoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_Fl_genes version 1.0
; SEQ ID NO 69
; LENGTH: 1690
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128)..(1522)
US-09-620-312D-69

Query Match 2.1%; Score 42.8; DB 4; Length 1690;
Best Local Similarity 47.9%; Pred. No. 0.013;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;

QY 1016 CCGAGAACTTCAAGAGCCACCGATGAGCTGAGCTCAAGCTCGCCCAAGCTGAGTGAT 1075

DB 55 CAAGGAGTTGCAACCCAGGTGGCGCAGACCTGGACGACGAGCTGGCATGGTTCCAGAGCG 114
QY 1076 CAAGGGA---TCCTGGCAGCCCGTGGGGATCTCTCTATTGACTCTCTCCAGATCACCT 1132
DB 115 GCTGCCACTGTGCCATGACAGACGAGCGAGGCAACGGTTTGAGCGCGTCCAGACGACAT 174
QY 1133 CGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCCCTCTCTGAAAGAGAACCTGAGCCACGT 1192
DB 175 CAAAGAAACACAGGCGCTGGCGGGAGATCCAGGCGCATGGCGCGCGCTGGAGGAGT 234
QY 1193 CAATGACCTTCTGCGCAGCTTACCACTTTGGGCAATTCAGCTCTCAGCGTATACCTCAG 1252
DB 235 GCTGAGCGCGCGCGCGCTGGCGTCTGCTGCGCAGCCCGAGGACAGAGCAGTGCGCCG 294
QY 1253 CACTCTGGAACACCTCTGAACACCACTGGAAGCTTCTGAGGTGGCCCTTCGAGGACCCGAGT 1312
DB 295 GGGCTGGAGCAGCTGCAGAGCGCTGGCGCGGACTGCGGAGGCTGCCGAGCGCAGCGCA 354
QY 1313 CAGGAGCTGCATGAAGCCACAGG 1338
DB 355 GCAGGTGCTGAGCGCGCTTCCAGG 380

RESULT 15
US-09-368-590-1
; Sequence 1, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; TITLE OF INVENTION: AUTOANTIGENS OF AUTOIMMUNE DISEASES
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; CURRENT FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7812
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(6879)
; NAME/KEY: unsure
; LOCATION: (100)...(102)
; NAME/KEY: unsure
; LOCATION: (1021)...(1023)
; NAME/KEY: unsure
; LOCATION: (2266)...(2268)
US-09-368-590-1

Query Match 2.1%; Score 42.8; DB 3; Length 7812;
Best Local Similarity 47.9%; Pred. No. 0.037;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;

QY 1016 CCGAGAACTTCAAGAGCCACCGATGAGCTGAGCTCAAGCTCGCCCAAGCTGAGTGAT 1075
DB 3639 CAAGGAGTTGACCAAGCGGTGGCGCAGACCTTGAGCAGAGCTGGCATGGTTTCAGGAGCG 3698
QY 1076 CAAGGGA---TCCTGGCAGCCCGTGGCGATCTCTCTATTGACTCTCTCCAAAGATCACCT 1132
DB 3699 GCTGCCACTGGCCATGACAGACGAGCGAGGCAACGGTTTGACGGCGTCCAGACGACAT 3758
QY 1133 CGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCCCTCTCTGAAAGAGAACCTGAGCCACGT 1192
DB 3759 CAAAGAAACACAGGCGCTTCGCGCGGAGATCCAGCGCATGGCGCGCGCTTCGAGGAGT 3818
QY 1193 CAATGACCTTCTGCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAACCTCAG 1252
DB 3819 GCTGAGCGCGCGCGCTGGCGTCTCTGCGCAGCCCGGAGGAGGAGTGCGCCG 3878

QY	1253	CACTCTGGAAGACTGAACACCAAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGT	1312
Db	3879	GGGCTTGGACGACGCTGCAGAGCCCTTGGGCCGACTGCCGGAGGCTGCCGAGCGACGGCA	3938
QY	1313	CAGGCACTGCATGAAGCCACAGGG	1338
Db	3939	GCAGGTGCTGGACGCGCCTTCCAGG	3964

Search completed: September 19, 2004, 11:07:02
 Job time : 106.19 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 19, 2004, 08:30:27 ; Search time 666.733 Seconds
(without alignments)
15148.395 Million cell updates/sec

Title: US-09-845-416-14_COPY_1000_3000

Perfect score: 2001
Sequence: 1 ggcagttcattgatggagag.....cattgagcgaagtgtccgga 2001

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3327077 seqs, 253723180 residues

Total number of hits satisfying chosen parameters: 6654154

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq.*
- 18: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq.*
- 19: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2001	100.0	3446	10	US-09-845-416-14 Sequence 14, Appl
2	1990	99.5	4414	10	US-09-845-416-32 Sequence 32, Appl
3	1820.8	91.0	5417	16	US-10-149-736-39 Sequence 39, Appl
4	1593	79.6	3510	10	US-09-845-416-12 Sequence 12, Appl
5	1593	79.6	4476	10	US-09-845-416-31 Sequence 31, Appl
6	1263	63.1	3858	10	US-09-845-416-9 Sequence 9, Appl
7	1263	63.1	4825	10	US-09-845-416-29 Sequence 29, Appl
8	1263	63.1	4848	10	US-09-845-416-35 Sequence 35, Appl
9	1263	63.1	5060	10	US-09-845-416-36 Sequence 36, Appl
10	1254	62.7	5339	16	US-10-149-736-40 Sequence 40, Appl
11	1251.2	62.5	3531	10	US-09-845-416-10 Sequence 10, Appl
12	1251.2	62.5	4498	10	US-09-845-416-30 Sequence 30, Appl
13	1245	62.2	4182	10	US-09-845-416-2 Sequence 2, Appl
14	1245	62.2	5149	10	US-09-845-416-27 Sequence 27, Appl

15	1132.2	56.6	5462	16	US-10-149-736-41	Sequence 41, Appl
16	1112	55.6	3999	10	US-09-845-416-6	Sequence 6, Appl
17	1112	55.6	4966	10	US-09-845-416-28	Sequence 28, Appl
18	1112	55.6	4990	10	US-09-845-416-34	Sequence 34, Appl
19	1002.6	50.1	1821	10	US-09-845-416-13	Sequence 13, Appl
20	1002.6	50.1	2169	10	US-09-845-416-4	Sequence 4, Appl
21	1002.6	50.1	8689	16	US-10-149-736-42	Sequence 42, Appl
22	1002.6	50.1	11058	10	US-09-845-416-1	Sequence 1, Appl
23	1002.6	50.1	11443	16	US-10-149-736-44	Sequence 44, Appl
24	1002.6	50.1	12057	16	US-10-149-736-47	Sequence 47, Appl
25	1002.6	50.1	13957	9	US-09-782-378A-22	Sequence 22, Appl
26	1002.6	50.1	13957	9	US-09-880-107-2284	Sequence 2284, Ap
27	1002.6	50.1	13957	16	US-10-149-736-1	Sequence 1, Appl
28	1002.6	50.1	14069	13	US-10-342-887-434	Sequence 434, App
29	1002.6	50.1	14069	13	US-10-342-887-434	Sequence 434, App
30	1002.6	50.1	14082	13	US-10-172-118-434	Sequence 981, App
31	1002.6	50.1	14082	13	US-10-342-887-981	Sequence 981, App
32	1002.6	50.1	14082	16	US-10-172-118-981	Sequence 108, App
33	1001	50.0	1434	10	US-10-341-434-108	Sequence 15, Appl
34	985.8	49.3	1391	10	US-09-845-416-3	Sequence 3, Appl
35	869.8	43.5	13815	16	US-10-149-736-2	Sequence 2, Appl
36	665	33.2	1667	10	US-09-845-416-7	Sequence 7, Appl
37	514.6	25.7	10705	12	US-10-152-319A-1598	Sequence 1598, Ap
38	506.6	25.3	11096	16	US-10-149-736-4	Sequence 23, Appl
39	490.6	24.5	10302	9	US-09-782-378A-23	Sequence 3, Appl
40	490.6	24.5	10302	16	US-10-149-736-3	Sequence 667, App
41	481.8	24.1	16531	15	US-10-101-510-667	Sequence 157, App
42	470.2	23.5	5106	13	US-10-220-120-157	Sequence 35, Appl
43	457	22.8	887	16	US-10-149-736-35	Sequence 11, Appl
44	350	17.5	1340	10	US-09-845-416-11	Sequence 8, Appl
45	327	16.3	327	16	US-10-149-736-8	

ALIGNMENTS

RESULT 1

US-09-845-416-14
; Sequence 14, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 3446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-14

Query Match	100.0%;	Score 2001;	DB 10;	Length 3446;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2001;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GGCAGTTCAATTGATGGAGAGTAAAGTAAACCTGACCGTTATCAACAGCTTTAGAGAA	60	
Db	1000	GGCAGTTCAATTGATGGAGAGTAAAGTAAACCTGACCGTTATCAACAGCTTTAGAGAA	1059	
Qy	61	GTATTATCGTGGCTTTCTTCTGCTGAGGACACATTCGAGGAGAGATTCTTAAT	120	
Db	1060	GTATTATCGTGGCTTTCTTCTGCTGAGGACACATTCGAGGAGAGATTCTTAAT	1119	
Qy	121	GATGGAAGTGGTGAAGACACAGCTTTCATCTCATGAGGGGTACATGATGATTGACA	180	
Db	1120	GATGGAAGTGGTGAAGACACAGCTTTCATCTCATGAGGGGTACATGATGATTGACA	1179	

181 G C C C A T C A G G G C C G G T T G T A T A T T C T A C A A T T G G G A A G T A A G C T G A T T G G A A C A G A 240
1180 G C C C A T C A G G G C C G G T T G T A T A T T C T A C A A T T G G G A A G T A A G C T G A T T G G A A C A G A 1239
241 A A A T T A T C A G A A G A T G A A G A A C T G A A G T A C A A G A G C A G A T A A T C T C T A A A T T C A A G A 300
1240 A A A T T A T C A G A A G A T G A A G A A C T G A A G T A C A A G A G C A G A T A A T C T C T A A T T C A A G A 1299
301 T G G A A T G C C T C A G G T A G C T A G C A T G G A A A A C A A A G C A A T T A C A T A G A G T T T A A T G 360
1300 T G G A A T G C C T C A G G T A G C T A G C A T G G A A A A C A A A G C A A T T T A C A T A G A G T T T A A T G 1359
361 G A T C T C C A G A T C G A A A C T G A A G A G T T G A A T G A C T G G C T A C A A A A C A G A A G A A G A A 420
1360 G A T C T C C A G A T C G A A A C T G A A G A G T T G A A T G A C T G G C T A C A A A A C A G A A G A A G A A 1419
421 C A A G A A A T T G A G A G A G A G C C T C T T G G A C C T G A T C T T G A A G A C C T A A A C G C C A A G T A C 480
1420 C A A G A A A T T G A G A G A G A G C C T C T T G G A C C T G A T C T T G A A G A C C T A A A C G C C A A G T A C 1479
481 A A C A A C A T A A G T G C T T C A A G A A G A T C T A G A A C A A G A A C A A G T C A G G T C A A T T C T C T C A 540
1480 A A C A A C A T A A G T G C T T C A A G A A G A T C T A G A A C A A G A A C A A G T C A G G T C A A T T C T C T C A 1539
541 C T C A C A T G T G T G T A G T T G A T G A A T C T A G T G G A G A T C A C G C A A C T G C T G T T T G G A G 600
1540 C T C A C A T G T G T G T A G T T G A T G A A T C T A G T G G A G A T C A C G C A A C T G C T G T T T G G A G 1599
601 A A C A C T T A A G T A T T G G A G A T C G A T G G G C A A A C A T C T G T A G T G G A C A G A A G C C G C T 660
1600 A A C A C T T A A G T A T T G G A G A T C G A T G G G C A A A C A T C T G T A G T G G A C A G A A G C C G C T 1659
661 G G G T C T T T T A C A A G A C A T C C T T C T C A A T G S A C G C T T A C T G A A G A A C A G T G C C T T 720
1660 G G G T C T T T T A C A A G A C A T C C T T C T C A A T G S A C G C T T A C T G A A G A A C A G T G C C T T 1719
721 T T A G T G C A T G C T T T C A G A A A A G A A G A T G C A G T G A A C A A G A T T C A C A A C A T G C C T T T A 780
1720 T T A G T G C A T G C T T T C A G A A A A G A A G A T G C A G T G A A C A A G A T T C A C A A C A T G C C T T T A 1779
781 A A G A T C A A A A T G A A T T A T C A G T T C A A A A C T G C C C T T T T A A A G C G G A T C T A G 840
1780 A A G A T C A A A A T G A A T T A T C A G T T C A A A A C T G C C C T T T T A A A G C G G A T C T A G 1839
841 A A A G A A A A G C A A T C C A T G G G C A A A C T G T A T T C A C T C A A A C A A G A T C T T T T C A A C A C 900
1840 A A A G A A A A G C A A T C C A T G G G C A A A C T G T A T T C A C T C A A A C A A G A T C T T T T C A A C A C 1899
901 T G A A G A A T A A G T C A G T G A C C C A G A A C G G A A G C A T G G C T G G A T A A C T T T G C C C G T G T T 960
1900 T G A A G A A T A A G T C A G T G A C C C A G A A C G G A A G C A T G G C T G G A T A A C T T T G C C C G T G T T 1959
961 G G G A T A A T T A G T C C A A A A A C T T G A A A G A G T A C A C A G A C C C T T G A A A G A C T C C A G G 1020
1960 G G G A T A A T T A G T C C A A A A A C T T G A A A G A G T A C A C A G A C C C T T G A A A G A C T C C A G G 2019
1021 A A C T T C A A G A G C C A C G G A T A G C T G G A C C T C A A G T G G C C A A G C T G A G G T G A T C A A G 1080
2020 A A C T T C A A G A G C C A C G G A T A G C T G G A C C T C A A G T G G C C A A G C T G A G G T G A T C A A G 2079
1081 G A T C C T G G A C C G G T G G G G A T C C T C A T T G A C T C T C C A A G A T A C C T C G A G A A G 1140
2080 G A T C C T G G A C C G G T G G G G A T C C T C A T T G A C T C T C C A A G A T A C C T C G A G A A G 2139
1141 T C A A G G C A C T T C G A G G A G A A A T G C C C T C T G A A G A A A C G T G A G C C A C G T C A A T G A C C 1200
2140 T C A A G G C A C T T C G A G G A G A A A T G C C C T C T G A A G A A A C G T G A G C C A C G T C A A T G A C C 2199
1201 T T G C T G C C A G C T A C A C T T T G G G A T T C A G T C T C A C G T A T A C C T C A G C A C T C T G G 1260
2200 T T G C T G C C A G C T A C A C T T T G G G A T T C A G T C T C A C G T A T A C C T C A G C A C T C T G G 2259
1261 A A G A C C T G A A C A C C A G A T G G A A G C T T C T G C A G G T G C C G C T G C A G G A C C G A G T C A G G C A G C 1320

2260 A A G A C C T G A A C A C A C A G A T G G A A G C T T C T G C A G T G C C G T C G A G G A C C G A G T C A G C A G C 2319
1321 T G C A T G A A G C C C A C A G G A C T T T G G T C C A G C A T C T C A G C A C T T T T T T C C A G C T G T G C C 1380
2320 T G C A T G A A G C C C A C A G G A C T T T G G T C C A G C A T C T C A G C A C T T T T T T C C A G C T G T G C C 2379
1381 A G G T C C C T G G A G A G A G C C A T C T C G C C A A A C A A A G T G C C T A C T A T A T C A A C C A G A 1440
2380 A G G T C C C T G G A G A G A G C C A T C T C G C C A A A C A A A G T G C C T A C T A T A T A C A A C C A G A 2439
1441 C T C A A A C A A C T T G C T G G A C C A C C C A A A A T G A C A G A G T C T A C C A G T C T T T A G C T G A C C 1500
2440 C T C A A A C A A C T T G C T G G A C C A C C C A A A A T G A C A G A G T C T A C C A G T C T T T A G C T G A C C 2499
1501 T G A A T A A T G T C A G A T T C T A G C T T A T A G A C T G C C A T G A A A C T C C G A A G A C T C C A A G G 1560
2500 T G A A T A A T G T C A G A T T C T A G C T T A T A G A C T G C C A T G A A A C T C C G A A G A C T C C A A G G 2559
1561 C C C T T T G C T T G G A T C T C T G A G C C T G T C A G C T G C A T G T G A T G C T T G G A C C A C A C A A C C 1620
2560 C C C T T T G C T T G G A T C T C T G A G C C T G T C A G C T G C A T G T G A T G C T T G G A C C A C A C A A C C 2619
1621 T C A A G A A A A T G A C C A G C C C A T G A T A T C C T G C A G A T T A T A A T T G T T G A C C A C A T A T T 1680
2620 T C A A G A A A A T G A C C A G C C C A T G A T A T C C T G C A G A T T A T A A T T G T T G A C C A C A T A T T 2679
1681 A T G A C C C C T G G A G A G A G A C A C A A A T T T G G T C A A G C T C C C T C T C G T G T G A T A T G T 1740
2680 A T G A C C C C T G G A G A G A G A C A C A A A T T T G T C A A C G T C C C T C T C T G T G G A T A T G T 2739
1741 G T C T G A A C T G G C T G C T G A A T G T T A T G A T A C G G G A C G A A C A G G A G A T C C G T G C C T G T 1800
2740 G T C T G A A C T G G C T G C T G A A T G T T A T G A T A C G G G A C G A A C A G G A G A T C C G T G C C T G T 2799
1801 C T T T T A A A C T G G A C A T A T T C C C T G T A A A C A C A T T T G G A A G A C A A G T A C A G A T A C C 1860
2800 C T T T T A A A C T G G A C A T A T T C C C T G T A A A C A C A T T T G G A A G A C A A G T A C A G A T A C C 2859
1861 T T T T C A A G C A A G T G G C A A G T T C A A C A G A T T T T G T G A C C A G C C A G C T G G G C C T C C T T C 1920
2860 T T T T C A A G C A A G T G G C A A G T T C A A C A G A T T T T G T G A C C A G C C A G C T G G G C C T C C T T C 2919
1921 T G C A G A T T C T A T C C A A A A T T C C A A G A C A G T T G G G T G A A G T T G C A T C C T T T G G G G G C A G T A 1980
2920 T G C A G A T T C T A T C C A A A A T T C C A A G A C A G T T G G G T G A A G T T G C A T C C T T T G G G G G C A G T A 2979
1981 A C A T T G A G C C A A G T G T C C G G A 2001
2980 A C A T T G A G C C A A G T G T C C G G A 3000

RESULT 2
US-09-845-416-32
; Sequence 32, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 4414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-32

Query Match				99.5%;	Score 1990;	DB 10;	Length 4414;
Best Local Similarity				100.0%;	Pred. No. 0;		
Matches 2001; Conservative				0;	Mismatches	0;	Indels 1; Gaps 1;
QY	1	GGCAGTTCATTGATGAGAGTGAAGTAAACCTTGGACCGTTTATCAACACGCTTTAGAGAA	60				
DB	1757	GGCAGTTCATTGATGAGAGTGAAGTAAACCTTGGACCGTTTATCAACACGCTTTAGAGAA	1816				
QY	61	GTATTATCGTGGCTTCTTCTGCTGAGGACACANTTGAAGCAACAGAGAGATTTCTAAT	120				
DB	1817	GTATTATCGTGGCTTCTTCTGCTGAGGACACANTTGAAGCAACAGAGAGATTTCTAAT	1876				
QY	121	GATGTGGAAGTGTGGAAGACCAAGTTTCTATCTATGAGGGGTACATGATGATTTGACA	180				
DB	1877	GATGTGGAAGTGTGGAAGACCAAGTTTCTATCTATGAGGGGTACATGATGATTTGACA	1936				
QY	181	GCCCATCAGGCGGGTGTGTAATATTTCAATTTGGGAAGTAAGCTGATTTGACAACAGGA	240				
DB	1937	GCCCATCAGGCGGGTGTGTAATATTTCAATTTGGGAAGTAAGCTGATTTGACAACAGGA	1996				
QY	241	AAATTATCAGAGATCAAGAACTGAAGTACAGAGCAGATGATCTCTTAATTTCAAGA	300				
DB	1997	AAATTATCAGAGATCAAGAACTGAAGTACAGAGCAGATGATCTCTTAATTTCAAGA	2056				
QY	301	TGGGAATGCTCTCAGGCTAGCTAGCATGGAAGCAATTTACATAGAGTTTAAATG	360				
DB	2057	TGGGAATGCTCTCAGGCTAGCTAGCATGGAAGCAATTTACATAGAGTTTAAATG	2116				
QY	361	GATCTCCAGAACTC-GAAACTGAAAGAGTTGAATGATGCTGGTAAACAAACAGAGAAAGA	419				
DB	2117	GATCTCCAGAACTCAGAAAGTGAAGAGTTGAATGATGCTGGTAAACAAACAGAGAAAGA	2176				
QY	420	ACAAGGAAATGAGGAAGAGCTCTTGGACCTGATCTTTGAAGACCTTAAAGCCCAAGTA	479				
DB	2177	ACAAGGAAATGAGGAAGAGCTCTTGGACCTGATCTTTGAAGACCTTAAAGCCCAAGTA	2236				
QY	480	CAACAAATAGGTGCTTCAAGAAATCTAGAAAGAACTAGAGTCAAGTCAAGTCAATTTCTCTC	539				
DB	2237	CAACAAATAGGTGCTTCAAGAAATCTAGAAAGAACTAGAGTCAAGTCAAGTCAATTTCTCTC	2296				
QY	540	ACTCACATGCTGTGTAGTGTGATGATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA	599				
DB	2297	ACTCACATGCTGTGTAGTGTGATGATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA	2356				
QY	600	GAACAACTTAAGGTATTTGGAGATCGATGGCAACATCTGTAGATGGAAGAACCGC	659				
DB	2357	GAACAACTTAAGGTATTTGGAGATCGATGGCAACATCTGTAGATGGAAGAACCGC	2416				
QY	660	TGGGTTCTTTTACAGACATCTTCTCAATGCGCAACGCTTACTGGAAGAACAGTGCCTT	719				
DB	2417	TGGGTTCTTTTACAGACATCTTCTCAATGCGCAACGCTTACTGGAAGAACAGTGCCTT	2476				
QY	720	TTTAGTGCACTGGCTTTTCAAGAAAGAGATGCAAGTGAACAGATTCACCAACTGGCTTT	779				
DB	2477	TTTAGTGCACTGGCTTTTCAAGAAAGAGATGCAAGTGAACAGATTCACCAACTGGCTTT	2536				
QY	780	AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGGCGTTTAAAGCGGATCTA	839				
DB	2537	AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGGCGTTTAAAGCGGATCTA	2596				
QY	840	GAAAGAAAAAGCAATCCATGGCAACCTGATTTCACTCAAAACAGATCTTTTCAACA	899				
DB	2597	GAAAGAAAAAGCAATCCATGGCAACCTGATTTCACTCAAAACAGATCTTTTCAACA	2656				
QY	900	CTGAAGAAATAGTCAAGTACCCAGAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGT	959				
DB	2657	CTGAAGAAATAGTCAAGTACCCAGAGACGGAAGCATGGCTGGATAAATTTGCCCGGTGT	2716				
QY	960	TGGGATAATTTAGTCCAAAACCTTGAAGAGATGACGACAGACCTTGAAGACTCCAG	1019				
DB	2717	TGGGATAATTTAGTCCAAAACCTTGAAGAGATGACGACAGACCTTGAAGACTCCAG	2776				
QY	1020	GAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGGTGATCAAG	1079				

RESULT 3
US-10-149-736-39
; Sequence 39, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:

DB	2777	GAATCTTCAAGAGGCCACCGATGAGCTGAGCTCAAGCTGGCCCAAGTGAAGTATCAAG	2836
QY	1080	GGATCTCTGGCAGCCCGTGGCGGATCTCTCATTCAGCTCTCTCAGAGTCACTCGAGAA	1139
DB	2837	GGATCTCTGGCAGCCCGTGGCGGATCTCTCATTCAGCTCTCTCAGAGTCACTCGAGAA	2896
QY	1140	GTCAAGGACATTTGAGGAGAAATTTGCGCTCTGAAAGAGAACGTTGAGCCAGCTCAATGAC	1199
DB	2897	GTCAAGGACATTTGAGGAGAAATTTGCGCTCTGAAAGAGAACGTTGAGCCAGCTCAATGAC	2956
QY	1200	CTTGCTCGCCAGCTTACCATTTTGGGCAATTCAGCTCTCAGCGTATAACCTCAGCACTCTG	1259
DB	2957	CTTGCTCGCCAGCTTACCATTTTGGGCAATTCAGCTCTCAGCGTATAACCTCAGCACTCTG	3016
QY	1260	GAAACCTGAAACCAACATGGAAGCTTTGTCAGGTGCGCCGTCGAGGACGAGTCAGGAC	1319
DB	3017	GAAACCTGAAACCAACATGGAAGCTTTGTCAGGTGCGCCGTCGAGGACGAGTCAGGAC	3076
QY	1320	CTGCATGAAGCCACAGGAGCTTTTGGTCCAGCATCTCAGCACTTTCTTTTCCAGCTCTGTC	1379
DB	3077	CTGCATGAAGCCACAGGAGCTTTTGGTCCAGCATCTCAGCACTTTCTTTTCCAGCTCTGTC	3136
QY	1380	CAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAG	1439
DB	3137	CAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAG	3196
QY	1440	ACTCAAAACATTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGAC	1499
DB	3197	ACTCAAAACATTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGAC	3256
QY	1500	CTGAATATGTCAGATCTTCTAGGCTTATAGGACTGCCATGAACTCCGAGAGCTCGAAG	1559
DB	3257	CTGAATATGTCAGATCTTCTAGGCTTATAGGACTGCCATGAACTCCGAGAGCTCGAAG	3316
QY	1560	GCCCTTTGCTTGGATCTCTTGGAGCTGTCAGCTGTCATGTCCTTGGAGCCAGCAAC	1619
DB	3317	GCCCTTTGCTTGGATCTCTTGGAGCTGTCAGCTGTCATGTCCTTGGAGCCAGCAAC	3376
QY	1620	CTCAAGCAAAATGACAGCCATGATATCTGAGCTGTCATGTCCTTGGAGCCAGCAAC	1679
DB	3377	CTCAAGCAAAATGACAGCCATGATATCTGAGCTGTCATGTCCTTGGAGCCAGCAAC	3436
QY	1680	TATCAGCCCTGGAGCAAGAGCAACAAATTTGGTCAACGTCCTCTCTCGTGGATATG	1739
DB	3437	TATCAGCCCTGGAGCAAGAGCAACAAATTTGGTCAACGTCCTCTCTCGTGGATATG	3496
QY	1740	TGCTGAACTGGCTGCTGAATTTTATGATACGGGACGAAAGGAGGATCCGCTGCTG	1799
DB	3497	TGCTGAACTGGCTGCTGAATTTTATGATACGGGACGAAAGGAGGATCCGCTGCTG	3556
QY	1800	TCTTTTAAACCTGGCATCAATTTCCCTGCTGTAAGACCAATTTGGAAGCAAGTACAGATAC	1859
DB	3557	TCTTTTAAACCTGGCATCAATTTCCCTGCTGTAAGACCAATTTGGAAGCAAGTACAGATAC	3616
QY	1860	CTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTACCCAGCGCAGGCTGGGCTCCTT	1919
DB	3617	CTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTACCCAGCGCAGGCTGGGCTCCTT	3676
QY	1920	CTGCATGATCTATCCAAATTTCAAGACAGTTGGTGAAGTTCATCTCTTTGGGGGCGAGT	1979
DB	3677	CTGCATGATCTATCCAAATTTCAAGACAGTTGGTGAAGTTCATCTCTTTGGGGGCGAGT	3736
QY	1980	AACATTGAGCCAGTGTCCGGA 2001	
DB	3737	AACATTGAGCCAGTGTCCGGA 3758	

; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott O.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; PRIORITY FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 39
; LENGTH: 5417
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; US-10-149-736-39

Query Match 91.0%; Score 1820.8; DB 16; Length 5417;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 2; Indels 157; Gaps 2;

QY 1 GGCAGTTCATGTGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 60
Db 1199 GGCAGTTCATGTGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 1259

QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGCAAGCAAGAGGAGATTCTTAAT 120
Db 1259 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGCAAGCAAGAGGAGATTCTTAAT 1318

QY 121 GATGTGGAAGTGTGGAAGACCAAGTTTCATCTCATGAGGGTACATGATGATTGACA 180
Db 1319 GATGTGGAAGTGTGGAAGACCAAGTTTCATCTCATGAGGGTACATGATGATTGACA 1378

QY 181 GCCCATCAGGCGCGGTGGTAAATTTCTACAAATGGGAAGTAAAGTATTCGACAGGA 240
Db 1379 GCCCATCAGGCGCGGTGGTAAATTTCTACAAATGGGAAGTAAAGTATTCGACAGGA 1438

QY 241 AAATTATCAGAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCCTAAATTCAGAA 300
Db 1439 AAATTATCAGAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCCTAAATTCAGAA 1498

QY 301 TGGGAATGCCCTCAGGTAGCTAGCTAGGAAACAAAGCAATTTACATAGAGTTTAATG 360
Db 1499 TGGGAATGCCCTCAGGTAGCTAGCTAGGAAACAAAGCAATTTACATAGAGTTTAATG 1558

QY 361 GATCTCCAGAACTC-GAAATCTGAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGA 419
Db 1559 GATCTCCAGAACTCAGAACTGAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGA 1618

QY 420 ACAAGGAAATGAGGAAGAGCTCTTGGACCTGATCTTTGAAGACCTTAAAGCCCAAGTA 479
Db 1619 ACAAGGAAATGAGGAAGAGCTCTTGGACCTGATCTTTGAAGACCTTAAAGCCCAAGTA 1678

QY 480 CAACACATAGAGTGTCTCAAGAGATCTAGAACAGAACTAGTCAAGGTCAATTTCTCTC 539
Db 1679 CAACACATAGAGTGTCTCAAGAGATCTAGAACAGAACTAGTCAAGGTCAATTTCTCTC 1738

QY 540 ACTCATGCTGTGTGTGTGATGAAATCTAGTGGAGATCAGCAACTGCTCTTTGGAA 599
Db 1739 ACTCATGCTGTGTGTGTGATGAAATCTAGTGGAGATCAGCAACTGCTCTTTGGAA 1798

QY 600 GAACAACTTAAGGTATTCGAGATCGATGGGCAACATCTGTAGTGGACAGAGACGCG 659
Db 1799 GAACAACTTAAGGTATTCGAGATCGATGGGCAACATCTGTAGTGGACAGAGACGCG 1858

QY 660 TGGGTTCTTTTACAAGACATCTTCTCAATGGCAACGCTTTACTGAAGAAACAGTGCCTT 719
Db 1859 TGGGTTCTTTTACAAGACATCTTCTCAATGGCAACGCTTTACTGAAGAAACAGTGCCTT 1918

QY 720 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCGAGTGAACAGATTCCACAACTGGCTTT 779

Db 1919 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCGAGTGAACAGATTCCACAACTGGCTTT 1978

QY 780 AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAATGCGCGTTTAAAGCGGATCTA 839

Db 1979 AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAATGCGCGTTTAAAGCGGATCTA 2038

QY 840 GAAAAGAAAAAGCAATCCATGGCAAACTGTATTCTCACTCAAAACAAGATCTTCTTCAACA 899

Db 2039 GAAAAGAAAAAGCAATCCATGGCAAACTGTATTCTCACTCAAAACAAGATCTTCTTCAACA 2098

QY 900 CTGAAGATTAAGTCAAGTACCGACCCAGAGAGCGGAGCATGCTGGATAACTTTGCCCGGTCT 959

Db 2099 CTGAAGATTAAGTCAAGTACCGACCCAGAGAGCGGAGCATGCTGGATAACTTTGCCCGGTCT 2158

QY 960 TGGGATAATTTAGTCCAAAAAATTTGAAAAGAGTACAGCACAGA----- 1002

Db 2159 TGGGATAATTTAGTCCAAAAAATTTGAAAAGAGTACAGCACAGATTTTCACAGGCTGTCAAC 2218

QY 1003 ----- 1002

Db 2219 ACCACTCAGCCATCACTAACAACAGACAACTGTAAATGGAACAGTAACACTAGGTGAACACA 2278

QY 1003 ----- 1002

Db 2279 AGGAAACAGATCCTGTGTAAGCATGCTCAAGAGAACTTCCACCACCACTCCCCAAAAG 2338

QY 1003 -----CCCTTGAAAGACTCCAGAACTTCAAGAGGCCACGGATGAG 1043

Db 2339 AAGAGGCGAGATTACTGTGGATCTTGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAG 2398

QY 1044 CTGGACCTCAAGTCTGGCGCAAGCTGAGGTGATCAAGGGATCTCTGGCAGCCCGTGGCGCAT 1103

Db 2399 CTGGACCTCAAGTCTGGCGCAAGCTGAGGTGATCAAGGGATCTCTGGCAGCCCGTGGCGCAT 2458

QY 1104 CTCCTCATTCACCTCTCTCCAAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGAGAAATTT 1163

Db 2459 CTCCTCATTCACCTCTCTCCAAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGAGAAATTT 2518

QY 1164 GCGCTCTGAAAGAGAACTGAGCGACAGTCAATGACCTTGTCTGCCAGCTTACCACTTTG 1223

Db 2519 GCGCTCTGAAAGAGAACTGAGCGACAGTCAATGACCTTGTCTGCCAGCTTACCACTTTG 2578

QY 1224 GGCATTGAGCTCTCACCGTATAACCTCAGACCTCTGGAAGACCTGAACACAGATGGAAG 1283

Db 2579 GGCATTGAGCTCTCACCGTATAACCTCAGACCTCTGGAAGACCTGAACACAGATGGAAG 2638

QY 1284 CTTCTGAGGTGCGCGTCGAGGACCGAGTCAAGGAGTGCATGAAGCCCAAGGAGCTTT 1343

Db 2639 CTTCTGAGGTGCGCGTCGAGGACCGAGTCAAGGAGTGCATGAAGCCCAAGGAGCTTT 2698

QY 1344 GGTCCAGCATCTCAGCACTTTCTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGCCATC 1403

Db 2699 GGTCCAGCATCTCAGCACTTTCTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGCCATC 2758

QY 1404 TCGCCAAAAGAGTGGCCCTACTATATCAACACAGACTCAAAACAATTTGTCTGGGACCAT 1463

Db 2759 TCGCCAAAAGAGTGGCCCTACTATATCAACACAGACTCAAAACAATTTGTCTGGGACCAT 2818

QY 1464 CCGAAATGACAGCTCTACCACTGCTTGTAGCTGACCTGAATATGTCAGATTTCTCAGCT 1523

Db 2819 CCGAAATGACAGCTCTACCACTGCTTGTAGCTGACCTGAATATGTCAGATTTCTCAGCT 2878

QY 1524 TATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGC 1583

Db 2879 TATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTGCTTGGATCTCTTGAGC 2938

QY 1584 CTGTCAGCTGATGTGATGCTTTGGACCAACACCTCAAGCAAAATGACAGGCCCATG 1643

Db 2939 CTGTCAGCTGATGTGATGCTTTGGACCAACACCTCAAGCAAAATGACAGGCCCATG 2998

QY 1644 GATATCTGCGAGATTATTAATTTGTTGACCACTATTATGACCGCTTGGAGCAGAGCAC 1703

2999 GATATCCTCGAGATTATTAATGTTTGGACCACTATTATATGACCGCTCGAGCAAGACAC 3058
1704 AACAAATTTGGTCAACGTCCTCTCTCGGTGGATATGCTGCTGAACTGGCTGCTGAATGTT 1763
3059 AACAAATTTGGTCAACGTCCTCTCTCGGTGGATATGCTGCTGAACTGGCTGCTGAATGTT 3118
1764 TATGATACGGGACGAACAGGAGGATCCGTGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1823
3119 TATGATACGGGACGAACAGGAGGATCCGTGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3178
1824 CTCTGTAAGACCAATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCA 1883
3179 CTCTGTAAGACCAATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCA 3238
1884 ACAGGATTTTGTACACGAGCGAGCTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1943
3239 ACAGGATTTTGTACACGAGCGAGCTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3298
1944 AGACAGTTGGTGAAGTTGCATCCTTTTGGGGGAGTAAACATTCAGCAAGTTCGCGA 2001
3299 AGACAGTTGGTGAAGTTGCATCCTTTTGGGGGAGTAAACATTCAGCAAGTTCGCGA 3356

RESULT 4
US-09-845-416-12
; Sequence 12, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 12
; LENGTH: 3510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-12

Query Match 79.6%; Score 1593; DB 10; Length 3510;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 1811; Conservative 0; Mismatches 190; Indels 64; Gaps 4;
QY 1 GGCAGTTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAACACACTTTAGAAGAA 60
DB 1000 GGCAGTTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAACACACTTTAGAAGAA 1059
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAAT 120
DB 1060 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAAT 1119
QY 121 GATGTGAAGTGTGAAGACCAAGTTTCACTCATGAGGGGTACATGATGATTGACA 180
DB 1120 GATGTGAAGTGTGAAGACCAAGTTTCACTCATGAGGGGTACATGATGATTGACA 1179
QY 181 GCCCATCAGGCGCGGTGTGTAATATTCTACAAATTTGGGAAGTAAAGCTGATTTGGAACAGGA 240
DB 1180 GCCCATCAGGCGCGGTGTGTAATATTCTACAAATTTGGGAAGTAAAGCTGATTTGGAACAGGA 1239
QY 241 AAATTATCAAGAGTGAAGAACTGAAGTACAAGACAGATGATCTCTCTAAATTCAAGA 300
DB 1240 AAATTATCAAGAGTGAAGAACTGAAGTACAAGACAGATGATCTCTCTAAATTCAAGA 1299
QY 301 TGGGAATGCTCAGGAGTACATGGAAGAACTGAAGTACAAGACAGATGATCTCTCTAAATTCAAG 360
DB 1300 TGGGAATGCTCAGGAGTACATGGAAGAACTGAAGTACAAGACAGATGATCTCTCTAAATTCAAG 1359
QY 361 GATCTCCGAATC-GAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGA 419

DB 1360 GATCTCCAGAAATCAGAAATCGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGA 1419
QY 420 ACAAGGAAAAATGGAGGAGAGCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTA 479
DB 1420 ACAAGGAAAAATGGAGGAGAGCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTA 1479
QY 480 CAACAAATAGTGGTCTTCAGAGATCTAGAACAAAGCAAGTCAAGTCAAGTCAATCTCTCTC 539
DB 1480 CAACAAATAGTGGTCTTCAGAGATCTAGAACAAAGCAAGTCAAGTCAATCTCTCTC 1539
QY 540 ACTCACATGTGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 599
DB 1540 ACTCACATGTGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 1599
QY 600 GAAACAATTAAGTATTGGGAGATCGATGGCAACAATCTGTAGATGGAAGAGACGC 659
DB 1600 GAAACAATTAAGTATTGGGAGATCGATGGGCAACAATCTGTAGATGGAAGAGACGC 1659
QY 660 TGGGTCTCTTTTACAAAGACATCCTTTCTCAAAATGGCAAGCTCTTACTGAAGAACAGTGCCTT 719
DB 1660 TGGGTCTCTTTTACAAAGACATCTGACAGTGGAGAGCTCTGCACTTCTCTGCAGAA 1719
QY 720 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCAAGTGAACAGATTCACAACTGGCTTT 779
DB 1720 CTTCTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGC 1779
QY 780 AAAGATCAAAATGAAATGTTTATCAAGTCTTCARAACTGGCCGTTTTTAAAGCGGATCTA 839
DB 1780 GACTTTCCAGCAGTTTCAGAGCAACAGATGATACATAGGCGCTTCAGAGGAAATGAAA 1839
QY 840 GAAAAGAAAAAGCAATCCATGGGCAAACTGTA-----TTCACTCAAAACAAG 885
DB 1840 ACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAG 1899
QY 886 ATCTTCTTTCAACTGAGATTAAGTCAAGTACCCAGAGACCGGAGAGATGGTGA-- 943
DB 1900 CTTTGGAGGACTTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTCTGAGGAGAGA 1959
QY 944 -----TAACTTTGGCCCGGTGTGGGATAAATTTAGTCCAAA 978
DB 1960 GCCCAGAAATGTCACCTCGGCTTCTACGAAAGCAGCTGAGGAGGTCAATCTAGTGGGAA 2019
QY 979 AACTTGA-----AAGGTACAGACAGAGACCTTTGAAGACTC 1016
DB 2020 AAATTTGAACCTGCACTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAGACTC 2079
QY 1017 CAGGAATCTTCAAGAGGCCACGATGAGTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATC 1076
DB 2080 CAGGAATCTTCAAGAGGCCACGATGAGTGGACCTCAAGCTGCGCCCAAGCTGAGGTGATC 2139
QY 1077 AAGGATCTTGGCAGCCCGTGGGCGATCTCTCATTTGACTCTCTTCAAGATCACTCGAG 1136
DB 2140 AAGGATCTTGGCAGCCCGTGGGCGATCTCTCATTTGACTCTCTTCAAGATCACTCGAG 2199
QY 1137 AAGTCAAGGACATTCGAGGAGAAATTCGCTCTGAAAGAGACGTCGAGCCAGTCAAT 1196
DB 2200 AAGTCAAGGACATTCGAGGAGAAATTCGCTCTGAAAGAGACGTCGAGCCAGTCAAT 2259
QY 1197 GACCTTGTCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACGCTATAACCTCAGCACT 1256
DB 2260 GACCTTGTCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACGCTATAACCTCAGCACT 2319
QY 1257 CTGGAAGACCTTGAAACACCAAGTTCGAGGTGGCCCTCGAGGACCGGATCAGG 1316
DB 2320 CTGGAAGACCTTGAAACACCAAGTTCGAGGTGGCCCTCGAGGACCGGATCAGG 2379
QY 1317 CAGTGTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACCTCT 1376
DB 2380 CAGTGTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACCTCT 2439
QY 1377 GTCCAGGCTCCTCGGAGAGAGCCATCTGCGCAAAAGAGTGCCTCTACTATATCAACAC 1436

```
Db      2440 GTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCAC 2499
Qy      1437 GAGACTCAAACTTGTCTGGGACCAATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCT 1496
Db      2500 GAGACTCAAACTTGTCTGGGACCAATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCT 2559
Qy      1497 GACTGAATTAATGACAGTCTCAGCTTATAGAGTGCCTGCAATGAACTCCGAAAGACTGCAG 1556
Db      2560 GACTGAATTAATGACAGTCTCAGCTTATAGAGTGCCTGCAATGAACTCCGAAAGACTGCAG 2619
Qy      1557 AAGGCCCTTTGCTTGTGATCTCTTGGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 1616
Db      2620 AAGGCCCTTTGCTTGTGATCTCTTGGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 2679
Qy      1617 AACTCAAGCAAAATGACAGCCCATGGATATCTCGCAGATTATTAATTTGTTTGACCACT 1676
Db      2680 AACTCAAGCAAAATGACAGCCCATGGATATCTCGCAGATTATTAATTTGTTTGACCACT 2739
Qy      1677 ATTATGACCGCTGGAGCAAGACAAACAAATTTGCTCAAGCTCCCTCTCTGCGTGGAT 1736
Db      2740 ATTATGACCGCTGGAGCAAGACAAACAAATTTGCTCAAGCTCCCTCTCTGCGTGGAT 2799
Qy      1737 ATGTGCTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1796
Db      2800 ATGTGCTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2859
Qy      1797 CTGTCTTTTAAACTGGCATCATTTCCCTGCTGTAAGACACATTTGGAAGACAAAGTACAGA 1856
Db      2860 CTGTCTTTTAAACTGGCATCATTTCCCTGCTGTAAGACACATTTGGAAGACAAAGTACAGA 2919
Qy      1857 TACCTTTTCAAGCAAGTGGCAAGTCAACAGGATTTTGTGACAGCGCAGGCTGGCCCTC 1916
Db      2920 TACCTTTTCAAGCAAGTGGCAAGTCAACAGGATTTTGTGACAGCGCAGGCTGGCCCTC 2979
Qy      1917 CTCTGCTGATCTATCCAAATTCAGACAGTGGGTGAAAGTTCATCTCTTTGGGGC 1976
Db      2980 CTCTGCTGATCTATCCAAATTCAGACAGTGGGTGAAAGTTCATCTCTTTGGGGC 3039
Qy      1977 AGTAACATTGAGCAAGTGTCCGA 2001
Db      3040 AGTAACATTGAGCAAGTGTCCGA 3064

RESULT 5
US-09-845-416-31
; Sequence 31, Application US/09845416
; Publication No. US2003017132A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 4476
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-31

Query Match      79.6%; Score 1593; DB 10; Length 4476;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 1811; Conservative 0; Mismatches 190; Indels 64; Gaps 4;

Qy      1 GGCAGTTCATTGTGAGAGTGAAGTAAACCTGGACCGTTTATCAACAGCTTTAGAAAGAA 60
Db      1756 GGCAGTTCATTGTGAGAGTGAAGTAAACCTGGACCGTTTATCAACAGCTTTAGAAAGAA 1815
Qy      61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGCAAGACAGGAGAGATTCTTAAT 120
```

```
Db      1816 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTTAAT 1875
Qy      121 GATGTGAAGTGGTGAAGACCAAGTTTCATCTACTCATGAGGGGTACATGATGGATTGACA 180
Db      1876 GATGTGAAGTGGTGAAGACCAAGTTTCATCTACTCATGAGGGGTACATGATGGATTGACA 1935
Qy      181 GCCCATCAGGCCCGGGTGGTAAATATTCTACAATTTGGAAAGTAAAGCTGATTGGAACAGGA 240
Db      1936 GCCCATCAGGCCCGGGTGGTAAATATTCTACAATTTGGAAAGTAAAGCTGATTGGAACAGGA 1995
Qy      241 AAATTATCAAGAAGTAAAGAAACTGAAAGTACAAGACAGATGAATCTCTTAAATTTCAAGA 300
Db      1996 AAATTATCAAGAAGTAAAGAAACTGAAAGTACAAGACAGATGAATCTCTTAAATTTCAAGA 2055
Qy      301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGATTTTAATG 360
Db      2056 TGGGAATGCCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGATTTTAATG 2115
Qy      361 GATCTCCAGAAATC-GAAACTCAAGAGATTGAATGACTGGCTTAAACAAAGACAGAAAGA 419
Db      2116 GATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTTAAACAAAGACAGAAAGA 2175
Qy      420 ACAAGGAATGAGGAGAGAGCTCTTGGACCTGATCTTGAAGACCTTAAACCCCAAGTA 479
Db      2176 ACAAGGAATGAGGAGAGAGCTCTTGGACCTGATCTTGAAGACCTTAAACCCCAAGTA 2235
Qy      480 CAACAAATTAAGTGTCTTCAAGAAAGATCTAGAAACAAAGAACAAAGTCAGGTCATTTCTCTC 539
Db      2236 CAACAAATTAAGTGTCTTCAAGAAAGATCTAGAAACAAAGAACAAAGTCAGGTCATTTCTCTC 2295
Qy      540 ACTCAGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 599
Db      2296 ACTCAGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2355
Qy      600 GAACAACTTAAGTGTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCCGC 659
Db      2356 GAACAACTTAAGTGTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCCGC 2415
Qy      660 TGGTCTTTTCAAGACATCTTCTCAAATGCAAGCTCTTACTGAAGAACAGTGCCTT 719
Db      2416 TGGTCTTTTCAAGACATCTTCTCAAAGACAGTCTGACCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2475
Qy      720 TTTAGTGCATGGCTTTTCAAGAAAGAGATGAGTGAACAGATTCACAACTGGCTTT 779
Db      2476 CTTCTGTGTGGCTTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTTGAGGC 2535
Qy      780 AAGATCAAAATGAATGTTTATCAAGTCTTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTA 839
Db      2536 GACTTTCCAGCAGTTTCAAGAGCAGACGATGTACATAGGSCCTTCAAGAGGGAATTTGAA 2595
Qy      840 GAAAGAAAAGCAATCCATGGGCAACTGTA-----TTCACTCAAAACAAG 885
Db      2596 ACTAAGAAACCTGTAAATCATGAGTACTCTTGAAGATGTACGAATATTTTGAAGAGCAG 2655
Qy      886 ATCTTTCTTTCAACACTCAAGAATAAGTCAAGTACCCAGAGAGCGGAGCATGCTGGA-- 943
Db      2656 CTTTGGAGGACTAGAGAACTTACCAGAGGCCCAGAGAGCTGCTCTGAGGAGAGA 2715
Qy      944 -----TAACTTTCCCGGTGTTGGGTAATTTAGTCCAAA 978
Db      2716 GCCCAGAAATGCTACCTGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAA 2775
Qy      979 AACTTTGAA-----AAGATACAGACACAGACCCCTTGAAGACTC 1016
Db      2776 AAATTGAACCTGACTCCGCTGACTGGCAGAGAAATAATAGATGAGACCTTTGAAGACTC 2835
Qy      1017 CAGAACTTCAAGAGGCCACCGATGAGTGGACCTCAAGCTGCGCAAGCTGAGGTGATC 1076
Db      2836 CAGAACTTCAAGAGGCCACCGATGAGTGGACCTCAAGCTGCGCAAGCTGAGGTGATC 2895
Qy      1077 AAGGATCTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCAAGATCACCTCGAG 1136
```

Db 2896 AAGGATCTCGCAGCCCGTGGCGATCTCTCAATGACTCTCTCAAGATCACTCGAG 2955
QY 1137 AAAGTCAAGGCATCTCGAGGAGAAATTTGGCTCTGAAAGAGAAAGTGAAGCCAGCTCAAT 1196
Db 2956 AAAGTCAAGGCATCTCGAGGAGAAATTTGGCTCTGAAAGAGAAAGTGAAGCCAGCTCAAT 3015
QY 1197 GACCTTGTCTGGCAGCTTACCACTTTGGCATTTCAGCTCTCACCGTATTAACCTCAGCACT 1256
Db 3016 GACCTTGTCTGGCAGCTTACCACTTTGGCATTTCAGCTCTCACCGTATTAACCTCAGCACT 3075
QY 1257 CTGGAGAGCTGACACCAAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCAAG 1316
Db 3076 CTGGAAGAGCTGACACCAAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCAAG 3135
QY 1317 CAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCT 1376
Db 3136 CAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCT 3195
QY 1377 GTCCAGGTCCTGGAGAGAGCCATCTCGCCAAACAAAGTGGCTTACTATATCAACCAC 1436
Db 3196 GTCCAGGTCCTGGAGAGAGCCATCTCGCCAAACAAAGTGGCTTACTATATCAACCAC 3255
QY 1437 GAGACTCAAACTTGTGGGACCATCCCAAAATGACAGAGCTCTTACCAGTCTTTAGCT 1496
Db 3256 GAGACTCAAACTTGTGGGACCATCCCAAAATGACAGAGCTCTTACCAGTCTTTAGCT 3315
QY 1497 GACCTGAATAATGTCAGATCTCAGCTTATAGCATGCGCATGAACTCCGAAAGATGCGAG 1556
Db 3316 GACCTGAATAATGTCAGATCTCAGCTTATAGCATGCGCATGAACTCCGAAAGATGCGAG 3375
QY 1557 AAGGCCCTTGTCTGGATCTCTTGGCCTGTGAGTGTGATGCTGCTGAGCCATGAGCCAGC 1616
Db 3376 AAGGCCCTTGTCTGGATCTCTTGGCCTGTGAGTGTGATGCTGCTGAGCCATGAGCCAGC 3435
QY 1617 AACCTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTTAATTTGTTGACCACT 1676
Db 3436 AACCTCAAGCAAAATGACAGCCCATGGATATCTCTGAGATTTAATTTGTTGACCACT 3495
QY 1677 ATTATGACCGCTGAGAGAGACACAAATTTGGTCAACGTCCTCTCTCGCTGGAT 1736
Db 3496 ATTATGACCGCTGAGAGAGACACAAATTTGGTCAACGTCCTCTCTCGCTGGAT 3555
QY 1737 ATGTGTCTGAATCTGGCTGTGATGTTATGATACGGGACGAAACAGGAGGATCCGTGTC 1796
Db 3556 ATGTGTCTGAATCTGGCTGTGATGTTATGATACGGGACGAAACAGGAGGATCCGTGTC 3615
QY 1797 CTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGACAAGTACAGA 1856
Db 3616 CTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGACAAGTACAGA 3675
QY 1857 TACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACAGCGCAGGCTGGGCCTC 1916
Db 3676 TACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACAGCGCAGGCTGGGCCTC 3735
QY 1917 CTCTGATGATTTCTATCCAAATTCGAAGACAGTCTGGGTGAAGTTCATCTTTGGGGGC 1976
Db 3736 CTCTGATGATTTCTATCCAAATTCGAAGACAGTCTGGGTGAAGTTCATCTTTGGGGGC 3795
QY 1977 AGTAACATTCAGCAAGTGTCCGGA 2001
Db 3796 AGTAACATTCAGCAAGTGTCCGGA 3820

RESULT 6

US-09-845-416-9
; Sequence 9, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DB1142
; CURRENT APPLICATION NUMBER: US/09/845,416

; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3858
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-9

Query Match 63.18; Score 1263; DB 10; Length 3858;
Best Local Similarity 75.38; Pred. No. 0;
Matches 1816; Conservative 0; Mismatches 185; Indels 412; Gaps 3;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTTAAAGAA 60
Db 1000 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTTAAAGAA 1059
QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGCAAGCAGAGAGATTTCTAAT 120
Db 1060 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGCAAGCAGAGAGATTTCTAAT 1119
QY 121 GATGTGAAGTGTGAAAGACCACTTTCATCTCATGAGGGTACATGATGGATTTCACA 180
Db 1120 GATGTGAAGTGTGAAAGACCACTTTCATCTCATGAGGGTACATGATGGATTTCACA 1179
QY 181 GCCCATCAGGCCCGGTTGGTAAATTTCTCAATTTGGGAAGTAAAGCTGATTTGGAACAGGA 240
Db 1180 GCCCATCAGGCCCGGTTGGTAAATTTCTCAATTTGGGAAGTAAAGCTGATTTGGAACAGGA 1239
QY 241 AAATTTATCAGAGATGAAGAACTGAAGTCAAGAGCAGATGAATCTCTTAAATTTCAAGA 300
Db 1240 AAATTTATCAGAGATGAAGAACTGAAGTCAAGAGCAGATGAATCTCTTAAATTTCAAGA 1239
QY 301 TGGGAATTCCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTTACATAGTGTAAATG 360
Db 1300 TGGGAATTCCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTTACATAGTGTAAATG 1359
QY 361 GATCTCCAGATC-GAAACTGAAAGAGTTGAATGACTGGCTTAACAAACAGAGAAAGA 419
Db 1360 GATCTCCAGATCAGAACTGAAAGAGTTGAATGACTGGCTTAACAAACAGAGAAAGA 1419
QY 420 ACAAGGAAATGGAGGAGAGCCCTCTTGGACCTCATCTTGAAGACCTTAAACCGCCAAAGTA 479
Db 1420 ACAAGGAAATGGAGGAGAGCCCTCTTGGACCTCATCTTGAAGACCTTAAACCGCCAAAGTA 1479
QY 480 CAACAACTAAAGGTCTTCAAGAGATCTAGAACAGAAACAGTCAAGGTCAAGTCTCTC 539
Db 1480 CAACAACTAAAGGTCTTCAAGAGATCTAGAACAGAAACAGTCAAGGTCAAGTCTCTC 1539
QY 540 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCACTGCTGCTTTGGAA 599
Db 1540 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCACTGCTGCTTTGGAA 1599
QY 600 GAACAACTTAAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGACAGAGACCGC 659
Db 1600 GAACAACTTAAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGACAGAGACCGC 1659
QY 660 TGGGTTCTTTTACAAGACATCTTCTCAATGGCAAGCTCTTACTGAAGACAGTGCCTT 719
Db 1660 TGGGTTCTTTTACAAGACATCTAGATTTACTGCAAGTTCCTCCCTGGAGCTGGAAAAG 1719
QY 720 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCACTGAACAAAGATTCACAACTGGCTTT 779
Db 1720 TTTCTTGGCTTTCAGAGAGCTGAAACCACTGCCAATGTCTTACAGGATGCTACCGT 1779
QY 780 A-----AAGATCAAAATGAAATGTTATCAAGTCTTCAAAACTGGCGTT 824
Db 1780 AAGGAAAGGCTCTTAGAGAGCTCCAAAGGGAGTAAAGAGCTGTAGTGAACAACTGCAAGAC 1839
QY 825 TTAAAGGGATCTAGAAAAAGCAATCCATCCATGGCAACTGTATTCTACTCAAAACA 884

Dh 1840 CTCCAGGTGAATTTGAAGCTCACACAGATGTTTATCAAACTCGATGAAACAGCCAA 1899
Qy 885 GATCTCTTTCAACACTGAAGAATAAGTCAGTCAGCCAGGACGAGCAGCTGGTGGAT 944
Dh 1900 AAAATCTGAGATCCCTGGAGGTTCCGATGATGCGAGTCGTGTTACAAAGACGTTGGAT 1959
Qy 945 AACTTCCCGGTGTTGGGATAATTTAGTCCAAAACCTTGAAAGAGTACAGCAC-----999
Dh 1960 AACATGAATTCAGTGGAGTGAACCTTCGAAAAGTCTCTCAACATTAGTCCCATTTG 2019
Qy 1000 -----999
Dh 2020 GAAGCCAGTTCTGACCACTGGAAGCTCTGCACCTTTCTCTGAGGAACCTTCTGGTGG 2079
Qy 1000 -----999
Dh 2080 CTACAGCTGAAGATGATGAATTAAGCCGAGGACCTATTGGAGGCGACTTTCAGCA 2139
Qy 1000 -----999
Dh 2140 GTTCAGNAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAACTAAAGAACCT 2199
Qy 1000 -----999
Dh 2200 GTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCTTTGGAAGGA 2259
Qy 1000 -----999
Dh 2260 CTAGAGAACTTACCAGGAGCCAGAGAGTGCCTCTCTGAGAGAGAGCCAGAAATGTC 2319
Qy 1000 -----999
Dh 2320 ACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACCTAGTGGGAAATTTGAACCTG 2379
Qy 1000 -----AGACCTTTGAAGACTCCAGGAACCTTCAA 1028
Dh 2380 CAGCTCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTCAA 2439
Qy 1029 GAGGCCAGGATGAGTGGACCTCAAGCTGCGCAAGCTGAGTGTATCAAGGATCCTGG 1088
Dh 2440 GAGGCCAGGATGAGTGGACCTCAAGCTGCGCAAGCTGAGTGTATCAAGGATCCTGG 2499
Qy 1089 CAGCCCGTGGCGATCTCTCATATGACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCA 1148
Dh 2500 CAGCCCGTGGCGATCTCTCATATGACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCA 2559
Qy 1149 CTTGAGGAGAAATTCGGCTCTGAAAGAGACGTCAGCAGCTCAATGACCTTCTGCG 1208
Dh 2560 CTTGAGGAGAAATTCGGCTCTGAAAGAGACGTCAGCAGCTCAATGACCTTCTGCG 2619
Qy 1209 CAGCTTACCACTTTGGGCACTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTG 1268
Dh 2620 CAGCTTACCACTTTGGGCACTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTG 2679
Qy 1269 AACACAGATGGAAGCTTCTGAGTGGCGCTGAGGACCGAGTCAGGAGCTGATGAA 1328
Dh 2680 AACACAGATGGAAGCTTCTGAGTGGCGCTGAGGACCGAGTCAGGAGCTGATGAA 2739
Qy 1329 GCCACAGGAGCTTTGGTCCAGATCTCAGCACTTTCTTTCCACGCTCTGTCAGGAGTCC 1388
Dh 2740 GCCACAGGAGCTTTGGTCCAGATCTCAGCACTTTCTTTCCACGCTCTGTCAGGAGTCC 2799
Qy 1389 TGGGAGAGGCACTCTGCGCAAAAGTGCCTACTATATCAACAGAGACTCAAAACA 1448
Dh 2800 TGGGAGAGGCACTCTGCGCAAAAGTGCCTACTATATCAACAGAGACTCAAAACA 2859
Qy 1449 ACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGCTTTTACCTGACCTGAATAT 1508
Dh 2860 ACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGCTTTTACCTGACCTGAATAT 2919
Qy 1509 GTCAGATCTCAGCTTATAGGACTGCCATGAACCTCCGAAGACTCGAGAGGCCCTTTGC 1568
Dh 2920 GTCAGATCTCAGCTTATAGGACTGCCATGAACCTCCGAAGACTCGAGAGGCCCTTTGC 2979

RESULT 7

US-09-845-416-29
; Sequence 29, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 4825
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-29

Query Match 53.1%; Score 1263; DB 10; Length 4825;
Best Local Similarity 75.3%; Pred. No. 0;
Matches 1816; Conservative 0; Mismatches 185; Indels 412; Gaps 3;

Qy 1 GGCACTTCATGATGAGAGTGAAGTAAACCTGACCGTTATCAACAGCTTTAGAGAA 60
Dh 1757 GGCACTTCATGATGAGAGTGAAGTAAACCTGACCGTTATCAACAGCTTTAGAGAA 1816
Qy 61 GTATTATCGGGCTTTCTTCTGCTGAGGACACATTGCAAGCAGAGAGAGATTTCTAAT 120
Dh 1817 GTATTATCGGGCTTTCTTCTGCTGAGGACACATTGCAAGCAGAGAGATTTCTAAT 1876
Qy 121 GATGTGAAGTGTGAAGACCACTTCTACTCATGAGGGGTACATGATGATTTGACA 180
Dh 1877 GATGTGAAGTGTGAAGACCACTTCTACTCATGAGGGGTACATGATGATTTGACA 1936
Qy 181 GCCCATCAGGCGGGTGGTAAATATTCTACAATTGGGAAGTAAGCTGATGGAACAGGA 240

QY 1989 CCAAGTGTCCGGA 2001
Db 4157 CCAAGTGTCCGGA 4169

RESULT 8
US-09-845-416-35
; Sequence 35, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 4848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-35

Query Match 63.1%; Score 1263; DB 10; Length 4848;
Best Local Similarity 75.3%; Pred. No. 0;
Matches 1816; Conservative 0; Mismatches 185; Indels 412; Gaps 3;

QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTCGACCGTTATCAACACAGCTTTAGAAGAA 60
Db 1780 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTCGACCGTTATCAACACAGCTTTAGAAGAA 1839

QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGACCAAGGAGAGATTTCTAAT 120
Db 1840 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGACCAAGGAGAGATTTCTAAT 1899

QY 121 GATGTGAAGTGTGAAAGACCAAGTTTCATCTACTCATGAGGGTACATGATGATTGACA 180
Db 1900 GATGTGAAGTGTGAAAGACCAAGTTTCATCTACTCATGAGGGTACATGATGATTGACA 1959

QY 181 GCCCATCAGCGCGGTTGGTAATATCTTACAATGGGAAGTAAGCTGATTGGAACAGGA 240
Db 1960 GCCCATCAGCGCGGTTGGTAATATCTTACAATGGGAAGTAAGCTGATTGGAACAGGA 2019

QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGA 300
Db 2020 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGA 2079

QY 301 TGGGAATCGCTCAGGCTAGCTAGCTAGGAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 2080 TGGGAATCGCTCAGGCTAGCTAGCTAGGAAACAAAGCAATTTACATAGAGTTTAAATG 2139

QY 361 GATCTCCAGATC-GAAACTGAAGAGTTGAATGACTGCTTAACAAAACAGAAAGAGA 419
Db 2140 GATCTCCAGATCAGAACTGAAGAGTTGAATGACTGCTTAACAAAACAGAAAGAGA 2199

QY 420 ACAAGGAAATGGAGGAAGCGCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTA 479
Db 2200 ACAAGGAAATGGAGGAAGCGCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTA 2259

QY 480 CAACACATGAAGTGTCTGAAGAGATCTAGAACAGAAACAAGTCAAGGTCTAATCTCTC 539
Db 2260 CAACACATGAAGTGTCTGAAGAGATCTAGAACAGAAACAAGTCAAGGTCTAATCTCTC 2319

QY 540 ACTCACATGTTGGTGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTCTTTGGAA 599
Db 2320 ACTCACATGTTGGTGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTCTTTGGAA 2379

QY 600 GAACAACTTAAGTATTTGGAGATCGATGGGCAACATCTGTAGATGGACAGAAACCGC 659
Db 2380 GAACAACTTAAGTATTTGGAGATCGATGGGCAACATCTGTAGATGGACAGAAACCGC 2439

QY 660 TGGGTTCTTTTACAAGACATCCTTTCTCAATGGCAAGCTCTTTACTAGAACACAGTGCCTT 719
Db 2440 TGGGTTCTTTTACAAGACATCCTTTCTCAATGGCAAGCTCTTTACTAGAACACAGTGCCTT 2499

QY 720 TTTAGTGCATGGCTTTTCAAAAAGAGATGCAAGTGAACAGATTTCAACAACTGGCTTT 779
Db 2500 TTTCTTGCTGGCTTTACAGAGCTGAAACCACTCCCAATGTCTTACAGGATGTACCCGT 2559

QY 780 A-----AAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACCTGGCGTT 824
Db 2560 AAGGAAAGGCTCCTAGAGAGCTCCAGGGAGTAAAGAGCTGATGAACAATGCAAGAC 2619

QY 825 TTTAAAGCGGATCTAGAAAAGAAAAGCAATCCATGGGCAAACTGTATTCACTCAACAA 884
Db 2620 CTCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCAA 2679

QY 885 GATCTTCTTCAACACTGAAGAATAAGTCAAGTGAACCCAGAGACGGGAAGCATGCTGGAT 944
Db 2680 AAAATCTTGAGATCCCTGGAGGTTCCGATGATGAGTCTCTGTACAAAGACGTTGGAT 2739

QY 945 AACTTTGCCCGGTTTGGGATAATTTAGTCCAAAACCTTGAAAAGATGACAGCAC----- 999
Db 2740 AACATGAATCTCAAGTGGAGTGAACCTCGGAAAAAAGTCTCTCAACATTAGTCCCATTTG 2799

QY 1000 ----- 999
Db 2800 GAAGCCAGTTCTGACAGTGAAGCGTCTGCACCTTTCTCTGCAGGAACTTTCTGTGTGG 2859

QY 1000 ----- 999
Db 2860 CTACAGCTGAAAGATGATGAATTAAGCCGCGCAGCACCCTATTGGAGCGACTTTCCAGCA 2919

QY 1000 ----- 999
Db 2920 GTTCAGAAAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAAACCTTAAAGAACCT 2979

QY 1000 ----- 999
Db 2980 GTAATCATGACTCTCTTGAGACTGTGAGCAATTTCTGACAGCAGCAGCTTTGGAAGGA 3039

QY 1000 ----- 999
Db 3040 CTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTC 3099

QY 1000 ----- 999
Db 3100 ACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAAAAAATTGAACCTG 3159

QY 1000 -----AGACCTTGAAGACTTCCAGGAACCTCAA 1028
Db 3160 CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGACTTCCAGGAACCTCAA 3219

QY 1029 GAGCCACGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGTGTCAAGGATCTCTGG 1088
Db 3220 GAGCCACGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGTGTCAAGGATCTCTGG 3279

QY 1089 CAGCCCTGGGCGATCTCTCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCA 1148
Db 3280 CAGCCCTGGGCGATCTCTCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCA 3339

QY 1149 CTTTCAGAGAAATTTGGCCCTCTGAAGAGAAAGTGAAGCCACGTCAATGACCTTGTCTGC 1208
Db 3340 CTTTCAGAGAAATTTGGCCCTCTGAAGAGAAAGTGAAGCCACGTCAATGACCTTGTCTGC 3399

QY 1209 CAGCTTACCACTTTGGGCAATTCAGCTCTCAACCTGATAAACCTCAGCACTCTTGGAAAGACCTG 1268
Db 3400 CAGCTTACCACTTTGGGCAATTCAGCTCTCAACCTGATAAACCTCAGCACTCTTGGAAAGACCTG 3459

QY 1269 AACACAGATGGAAGCTTTCTGAGGTTGGCGTTCGAGGACCGAGTCAAGGAGCTGCATGAA 1328
Db 3460 AACACAGATGGAAGCTTTCTGAGGTTGGCGTTCGAGGACCGAGTCAAGGAGCTGCATGAA 3519

QY 1329 GCCACAGGACATTCGGTCCAGCATCTCAGCATCTTCTTTCCAGTCTCTCCAGGTCCTC 1388
Db GCCACAGGACATTCGGTCCAGCATCTCAGCATCTTCTTTCCAGTCTCTCCAGGTCCTC 3579
QY 1389 TGGGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACAGAGACTCAACA 1448
Db TGGGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACAGAGACTCAACA 3639
QY 1449 ACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAAT 1508
Db ACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAAT 3699
QY 1509 GTCCAGATCTCAGCTATAGGACTGCCATGAACCTCCGAAGACTGCAAGAGCCCTTTGC 1568
Db GTCCAGATCTCAGCTATAGGACTGCCATGAACCTCCGAAGACTGCAAGAGCCCTTTGC 3759
QY 1569 TTGGATCTCTTGGACCTGTGAGCTGTGATGCTTGGACCCAGACAACTCAAGCAA 1628
Db TTGGATCTCTTGGACCTGTGAGCTGTGATGCTTGGACCCAGACAACTCAAGCAA 3819
QY 1629 AATGACAGCCCATGATATCTGACAGATTAATTTGTTTGGACCACTATTTATGACCGC 1688
Db AATGACAGCCCATGATATCTGACAGATTAATTTGTTTGGACCACTATTTATGACCGC 3879
QY 1689 CTGGACAGAGACAACTTTGTTCAAGCTCTCTCTGCTGGATATGTTCTGAAC 1748
Db CTGGACAGAGACAACTTTGTTCAAGCTCTCTCTGCTGGATATGTTCTGAAC 3939
QY 1749 TGGCTGCTGAATTTATGATACGGACGAAACAGGAGGATCGGTCTCTCTTTTAA 1808
Db TGGCTGCTGAATTTATGATACGGACGAAACAGGAGGATCGGTCTCTCTTTTAA 3999
QY 1809 ACTGGCATCATTTCCCTGTGTAAGACACATTTGGAAGACAAAGTACAGATACCTTTCAAG 1868
Db ACTGGCATCATTTCCCTGTGTAAGACACATTTGGAAGACAAAGTACAGATACCTTTCAAG 4059
QY 1869 CAAAGTGGCAAGTTCAACAGGATTTGTGACAGCGAGGCTGGGCTCTCTGATGAT 1928
Db CAAAGTGGCAAGTTCAACAGGATTTGTGACAGCGAGGCTGGGCTCTCTGATGAT 4119
QY 1929 TCTATCCAAATTCGAAGACAGTTGGTGAAGTTGCAATCTTTGGGGGACAGTAACATTTGAG 1988
Db TCTATCCAAATTCGAAGACAGTTGGTGAAGTTGCAATCTTTGGGGGACAGTAACATTTGAG 4179
QY 1989 CCAAGTGTCCGGA 2001
Db CCAAGTGTCCGGA 4192

RESULT 9
US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication No. US2003017132A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DB1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 5060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-36

Query Match 63.1%; Score 1263; DB 10; Length 5060;
Best Local Similarity 75.3%; Pred. No. 0;
Matches 1816; Conservative 0; Mismatches 185; Indels 412; Gaps 3;

QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAAGAA 60
Db GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAAGAA 2051
QY 61 GTATATTCGTCGCTTCTTCTGCTGAGACACATTTGCAAGCAACAGGAGAGATTTCTAAT 120
Db GTATATTCGTCGCTTCTTCTGCTGAGACACATTTGCAAGCAACAGGAGAGATTTCTAAT 2111
QY 121 GATGTGAAGTGTGAAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTTTGACA 180
Db GATGTGAAGTGTGAAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTTTGACA 2171
QY 181 GCCCATCAGGCCCGGGTTGGTAATTTCTCAATTTGGGAAGTAAAGCTGATTTGGAACAGGA 240
Db GCCCATCAGGCCCGGGTTGGTAATTTCTCAATTTGGGAAGTAAAGCTGATTTGGAACAGGA 2231
QY 241 AAATTTATCAGAAAGATGAAGAACTGAAGTCAAGAGCAGATGAATCTCTTAAATTTCAAGA 300
Db AAATTTATCAGAAAGATGAAGAACTGAAGTCAAGAGCAGATGAATCTCTTAAATTTCAAGA 2291
QY 301 TGGGAATGCTTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTTACATAGAGTTTAAATG 360
Db TGGGAATGCTTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTTACATAGAGTTTAAATG 2351
QY 361 GATCTCCAGATC-GAAACTGAAAGAGTGAATGACTGGCTTACAAACACAGAGAAAGA 419
Db GATCTCCAGATC-GAAACTGAAAGAGTGAATGACTGGCTTACAAACACAGAGAAAGA 2411
QY 420 ACAAGGAAATGAGGAGAGAGCTCTTGGACCTCATCTTGAAGACCTTAAACCGCAAGTA 479
Db ACAAGGAAATGAGGAGAGAGCTCTTGGACCTCATCTTGAAGACCTTAAACCGCAAGTA 2471
QY 480 CAAACAATAAGTGTCTTCAAGAAAGATCTAGAACAAAGCAAGTCAAGTCAAGTCAATCTCTC 539
Db CAAACAATAAGTGTCTTCAAGAAAGATCTAGAACAAAGCAAGTCAAGTCAAGTCAATCTCTC 2531
QY 540 ACTCACATGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 599
Db ACTCACATGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2591
QY 600 GAAACAATAAGT 659
Db GAAACAATAAGT 2651
QY 660 TGGGTCTCTTTTACAGACATCTCTCAATGCGCAACGCTCTTACTGAAGAACACGTCCTT 719
Db TGGGTCTCTTTTACAGACATCTCTCAATGCGCAACGCTCTTACTGAAGAACACGTCCTT 2719
QY 720 TTTAGTGCATGCTTTTACAGAAAGAGAGTGCAGTGAACCAAGATTCACAACTGGCTTT 779
Db TTTAGTGCATGCTTTTACAGAAAGAGAGTGCAGTGAACCAAGATTCACAACTGGCTTT 2771
QY 780 A-----AAGATCAAAATGAATGTTATCAAGTCTTCAAAACTGCTTCAAAACTGGCGTT 824
Db A-----AAGATCAAAATGAATGTTATCAAGTCTTCAAAACTGCTTCAAAACTGGCGTT 2824
QY 2772 AAGGAAAGGCTCTTAGAAGACTCCAAAGGAGTAAAGAGCTGTATGAAACCAATGGCAAGAC 2831
Db AAGGAAAGGCTCTTAGAAGACTCCAAAGGAGTAAAGAGCTGTATGAAACCAATGGCAAGAC 884
QY 825 TTAAGGCGGATCTAGAAAAGCAATCCATGGCAAACTGTATTCATCTCAAAACA 884
Db TTAAGGCGGATCTAGAAAAGCAATCCATGGCAAACTGTATTCATCTCAAAACA 2891
QY 885 GATCTTCTTCAACACTGAAGAAATGAAGTCAAGTACCAAGAGAGCGGAGCAATGCTGTGAT 944
Db GATCTTCTTCAACACTGAAGAAATGAAGTCAAGTACCAAGAGAGCGGAGCAATGCTGTGAT 2951
QY 945 AACTTTGCCCCGGTGTGGGATAATTTAGTCCAAACCTTGAAGAGAGTACAGAC----- 999
Db AACTTTGCCCCGGTGTGGGATAATTTAGTCCAAACCTTGAAGAGAGTACAGAC----- 3011
QY 2952 AACATGAACCTCAAGTGGAGTGAACCTTCGGGAAAGAGTCTCTCAACATTTAGTCCCATTTG 3011
Db AACATGAACCTCAAGTGGAGTGAACCTTCGGGAAAGAGTCTCTCAACATTTAGTCCCATTTG 999
QY 1000 ----- 999
Db ----- 3012
QY 3012 GAAGCCAGTTCTGACCAGTGAAGCGTGTGACCTTTCTGTCAGGAACCTTCTGGTGTGG 3071
Db GAAGCCAGTTCTGACCAGTGAAGCGTGTGACCTTTCTGTCAGGAACCTTCTGGTGTGG 3071

QY 1000 ----- 999
Db 3072 CTACAGCTGAAGATGATGAATTAAGCGCGAGGACCTATTGGAGGCGACTTTCAGCA 3131
QY 1000 ----- 999
Db 3132 GTTCAGAGCAGACGATGACATAGGGCCCTTCAAGAGGGAATTGAAGAACTTAAGAACT 3191
QY 1000 ----- 999
Db 3192 GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTGGAAGGA 3251
QY 1000 ----- 999
Db 3252 CTAGAGAACTCTACAGAGAGCCGAGAGCTGCCTCTTGAGAGAGAGCCGAGAAATGTC 3311
QY 1000 ----- 999
Db 3312 ACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTAGTGGGAAATTTGACCTG 3371
QY 1000 -----AGACCTTTGAAGACTCCAGGAACCTTCAA 1028
Db 3372 CACTCGCTGACTGGCAGAGAGAAAATAGATGAGACCTTTGAAGAGCTCCAGGAACCTTCAA 3431
QY 1029 GAGGCCACGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCCTGG 1088
Db 3432 GAGGCCACGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCCTGG 3491
QY 1089 CAGCCCGTGGCGATCTCTCATTTGACTCTCTCAAGATCACTCGAGAAATCAAGCA 1148
Db 3492 CAGCCCGTGGCGATCTCTCATTTGACTCTCTCAAGATCACTCGAGAAATCAAGCA 3551
QY 1149 CTTGAGGAGAAATTCGGCTCTGAAAGAGAGCTGAGGAGAGCTCAATGACCTCTGCTGC 1208
Db 3552 CTTGAGGAGAAATTCGGCTCTGAAAGAGAGCTGAGGAGAGCTCAATGACCTCTGCTGC 3611
QY 1209 CAGCTTACCACTTTGGGCACTTCAAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTG 1268
Db 3612 CAGCTTACCACTTTGGGCACTTCAAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTG 3671
QY 1269 AACACAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCAGGAGCTGATGAA 1328
Db 3672 AACACAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCAGGAGCTGATGAA 3731
QY 1329 GCCCAGAGGACTTTGGTCAGCATCTCAGCACTTTCTTTCCACGCTCTGTCAGGCTGCC 1388
Db 3732 GCCCAGAGGACTTTGGTCAGCATCTCAGCACTTTCTTTCCACGCTCTGTCAGGCTGCC 3791
QY 1389 TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACAGAGACTCAAAACA 1448
Db 3792 TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACAGAGACTCAAAACA 3851
QY 1449 ACTTGCTGGAGACCATCCCAATATGACAGAGCTCTACAGCTCTTTAGCTGACCTGAATAT 1508
Db 3852 ACTTGCTGGAGACCATCCCAATATGACAGAGCTCTACAGCTCTTTAGCTGACCTGAATAT 3911
QY 1509 GTGAGATTCAGCTTATAGGACTGCCATGAACTCCGAAGACTCGAGAGGCCCTTTGC 1568
Db 3912 GTGAGATTCAGCTTATAGGACTGCCATGAACTCCGAAGACTCGAGAGGCCCTTTGC 3971
QY 1569 TTGGATCTCTTGAGCTGTGAGCTGATGATGCTGAGAGAGAGCTGAGAGAGAGCTCAAGCA 1628
Db 3972 TTGGATCTCTTGAGCTGTGAGCTGATGATGCTGAGAGAGAGCTGAGAGAGAGCTCAAGCA 4031
QY 1629 AATGACAGAGCCCATGATATCTGAGATTAATTTGTTGACCACTATTTATGACCGC 1688
Db 4032 AATGACAGAGCCCATGATATCTGAGATTAATTTGTTGACCACTATTTATGACCGC 4091
QY 1689 CTGAGCAGAGAGCACAACATTTGFTCAAGCTCCCTCTCTGCTGGATATGCTGAAC 1748
Db 4092 CTGAGCAGAGAGCACAACATTTGFTCAAGCTCCCTCTCTGCTGGATATGCTGAAC 4151
QY 1749 TGSCCTGCTGAATCTTTATGATACGGAGCAACAGGAGGATCCGCTGCTCTTTTAAA 1808

RESULT 10

US-10-149-736-40

; Sequence 40, Application US/10149736

; Publication No. US20030216332A1

; GENERAL INFORMATION:

; APPLICANT: Chamberlain, Jeffrey S.

; APPLICANT: Harper, Scott Q.

; TITLE OF INVENTION: Mini-Dyestrophin Nucleic Acids and Peptide Sequences

; FILE REFERENCE: UN-06968

; CURRENT APPLICATION NUMBER: US/10/149,736

; CURRENT FILING DATE: 2002-06-17

; PRIORITY APPLICATION NUMBER: PCT/US01/31126

; PRIORITY FILING DATE: 2001-10-04

; PRIORITY APPLICATION NUMBER: 60/238,848

; PRIORITY FILING DATE: 2000-10-06

; NUMBER OF SEQ ID NOS: 96

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 40

; LENGTH: 5339

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic

US-10-149-736-40

Query Match

Best Local Similarity 62.7%; Score 1254; DB 16; Length 5339;

Matches 1621; Conservative 0; Mismatches 380; Indels 79; Gaps 6;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 60
Db 1199 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 1258
QY 61 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTCGACAGCAAGAGAGATTTCTAAT 120
Db 1259 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTCGACAGCAAGAGAGATTTCTAAT 1318
QY 121 GATGTGGAAGTGGTGAAGACCCAGTTTTCATCTCATGAGGGTACATGATGGATTTGACA 180
Db 1319 GATGTGGAAGTGGTGAAGACCCAGTTTTCATCTCATGAGGGTACATGATGGATTTGACA 1378
QY 181 GCCCATCAGGCGCGGGTGGTAAATATTTCTACATTTGGGAAGTGAAGTGGTGAAGCAGA 240
Db 1379 GCCCATCAGGCGCGGGTGGTAAATATTTCTACATTTGGGAAGTGAAGTGGTGAAGCAGA 1438
QY 241 AATATTATCAGAAGTGAAGAACTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 300
Db 1439 AATATTATCAGAAGTGAAGAACTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1498
QY 301 TGGGAATGCTCAGGGTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAG 360
Db 1499 TGGGAATGCTCAGGGTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAG 1558


```
QY 361 GATCTCCAGAAATC-GAAACTGAAGAGTTGAATGACTGGCTTAACAAAACACAGAGAAAGA 419
Db 1559 CAACAGTTCCCTCGACCTGGAAAAGTTTCTTGGCTGTACAGAGCTGAACAACT 1618
QY 420 ACAGGAAATGAGGAGAGCTCTTGGACCTGATCTTGAAGACCTTAAACCCCAAGTA 479
Db 1619 GCCAATGCTCAGGATGCTACCCGTAAGAAAGGCTCCTAGAAGACTCCAAGGGAGTA 1678
QY 480 CAACAAATGAAGTGTCTCAAGAGATCTAGAAACAAGAACAGTCAAGGTCATTTCTCTC 539
Db 1679 AAGAGCTGATGAACAATGCGCAGAGCTCCAGGCTGAATGAAGCTCACAGATGTT 1738
QY 540 ACTCAGATGTTGGTGTAG-----TTGATGAATCTAGTGGAGATCAGCGAAT 587
Db 1739 TATCAACACCTGGATGAACACAGCCAAATAATCCTGAGATCCCTGGAAGGTTCCGATGAT 1798
QY 588 GCTGCTTTGGAAGAACAACTTAAG--TATTGGGAGATCGATGGGCAACATCTCTAGA 644
Db 1799 GCACTCTGTTACAGAGCTTTGGATAACATGAATCTCAAGTGGAGTGAATCTCGAAA 1858
QY 645 TGGACAGAGACCGCTGGGTCTTTTACAGAGCATCTTCTCAAAATGGAAGCTCTTACT 704
Db 1859 AAGTCTCTCAACATTAGGTCCCATTTTGAAGCCAGTTCTGACCAGTGAAGGCTCTGCAC 1918
QY 705 GAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGAAAAGAGATCGTGAACAAGATT 764
Db 1919 CTTTCTCTCAGGAACCTTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAG 1978
QY 765 CACACAACTGGCTTTTAAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAATCTGGCCGTT 824
Db 1979 GCACCTATTGGAGCGCATTTCCAGCAGTTTCAGAACGAAACGATGTACATAGGGCCTTC 2038
QY 825 TTAAGAGCGGATCTAGAAAAGAAAAGCAATCATCGGGCAAACTGTA-----871
Db 2039 AAGAGGGAATGTAAGAACTTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATA 2098
QY 872 -TTCACTCAACAGATCTTCTTCAACACTGAAGAAATTAAGTCAGTGACCCAGAGACGG 930
Db 2099 TTCTGACAGACAGCCTTTGGAAGACTAGAGAACTCTACAGAGAGCCCAAGAGCTG 2158
QY 931 AAGCATGGCTGA-----TAACTTTGGCCGGTGTGGG 963
Db 2159 CTTCTCTGAGGAGAGAGCCAGATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCT 2218
QY 964 ATAATTAGTCCAAAACCTGTA-----AAGATGACAGACAG 1001
Db 2219 AATACTGAGTGGAAAATTTGAACCTGCACTCGCTGATGCGCAGAGAAAATAGATGAG 2278
QY 1002 ACCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGACCTCAAGCTGCGC 1061
Db 2279 ACCCTTGAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGACCTCAAGCTGCGC 2338
QY 1062 CAAGCTGAGTGATCAAGGATCTCTGACGCCGCTGGGGATCTCTCATTTGACTCTCTC 1121
Db 2339 CAAGCTGAGTGATCAAGGATCTCTGACGCCGCTGGGGATCTCTCATTTGACTCTCTC 2398
QY 1122 CAAGATCACCTCGAGAAAGTCAAGGACTTCAAGAGGAGAAATTCGCGCTCTGAAAGAGAAC 1181
Db 2399 CAAGATCACCTCGAGAAAGTCAAGGACTTCAAGAGGAGAAATTCGCGCTCTGAAAGAGAAC 2458
QY 1182 GTGAGCCACCTCAATGACCTTGTGCGCAGCTTACCTTTGGCAATTCAGTCTCAACCG 1241
Db 2459 GTGAGCCACCTCAATGACCTTGTGCGCAGCTTACCTTTGGCAATTCAGTCTCAACCG 2518
QY 1242 TATACTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGAGGTGGCCGTC 1301
Db 2519 TATACTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGAGGTGGCCGTC 2578
QY 1302 GAGGACCGAGTCAAGGAGCTGATGAAGCCCAAGGGACTTTGGTCCAGCATCTCAGCAC 1361
Db 2579 GAGGACCGAGTCAAGGAGCTGATGAAGCCCAAGGGACTTTGGTCCAGCATCTCAGCAC 2638
QY 1362 TTCTCTCCACGCTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCC 1421
```

RESULT 11

```
US-09-845-416-10
; Sequence 10, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 10
; LENGTH: 3531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-10
```

Query Match 62.5%; Score 1251.2; DB 10; Length 3531;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1623; Conservative 0; Mismatches 378; Indels 85; Gaps 6;

QY 1 GCCAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGCTTATCAACAGCTTTTAAAGAA 60
Db 1000 GCCAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGCTTATCAACAGCTTTTAAAGAA 1059

QY 61 GTATTATCGTGGCTTTCTTCTGCTGAGGACACATTCGACGACAGGAGAGATTTCTAAT 120
Db 1060 GTATTATCGTGGCTTTCTTCTGCTGAGGACACATTCGACGACAGGAGAGATTTCTAAT 1119
QY 121 GATGTGGAAGTGGTGAAGACCACTTTTATCTATCTATGAGGGGTACATGATGATTTGACA 180
Db 1120 GATGTGGAAGTGGTGAAGACCACTTTTATCTATGAGGGGTACATGATGATTTGACA 1179
QY 181 GCCCATCAGGGCGGGTGGTATATTTCTAATTTGGGAAGTAAGCTGATTTGGAACAGGA 240
Db 1180 GCCCATCAGGGCGGGTGGTATATTTCTAATTTGGGAAGTAAGCTGATTTGGAACAGGA 1239
QY 241 AAATTTATCAGAAGATGAAGAACTGAAGTACAGGACAGATGAATCTCTTAATTTCAAGA 300
Db 1240 AAATTTATCAGAAGATGAAGAACTGAAGTACAGGACAGATGAATCTCTTAATTTCAAGA 1299
QY 301 TGGGAATGCTCAGGTAGCTAGCATGGAAGAAACAAAGCAATTTATCATAGATTTTAATG 360
Db 1300 TGGGAATGCTCAGGTAGCTAGCATGGAAGAAACAAAGCAATTTATCATAGAACTCATAGA 1359
QY 361 GATCTCCAGAA-----TCGAACCTGAAGAGTTCGAATCACTGGCTACCAAAACAGAA 413
Db 1360 TTAGTGAACAGTTCCTCCCTGGACCTGGAAAGTTTCTTGGCTGGCTTACAGAACTGAA 1419
QY 414 GAAAGAACAAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGC 473
Db 1420 ACAAATGCCAATGCTACAGGATGCTACCCGTAAGGAAAGGCTCTAGAGACTCCAAAG 1479
QY 474 CAAGTACAAACAATAGGTGCTTCAAGAGATCTAGAACAGAAACAAGTCAGGTCAAT 533
Db 1480 GGAATAAAGAGCTGATGAACAAATGGCAAGACCTCCAAAGTGAATTTGAAGCTCACACA 1539
QY 534 TCTCTCACTCACATGTTGGTGGTGGTGGTATGATGA-----ATCTAGTGGAGATCAC 581
Db 1540 GATGTTTATCAACCTGGATGAACAGCCAAAGAAATCCTGATCCCTGGAAGTTCC 1599
QY 582 GCAACTGCTGTTTGAAGAACAACTTAAG-----TATTGGGAGATCGATGGGCAACATC 638
Db 1600 GATGATGAGTCTGTTTACAAAGACGTTTGGATAACATGAATTTCAAGTGGAGTGAACIT 1659
QY 639 TGTAGATGGACAGACGCTGGTCTTTTACAGACATCCTTCTCAATTTGGCAAGCT 698
Db 1660 CGGAAGAAAGTCTCTCAACATAGTCCATTTTGAAGCCAGTTCTGACCAAGTGGAGGCT 1719
QY 699 CTCTACTGAAGAACAGTGCCTTTTATGTCATGCTTTTGAAGAAAGAGATGAGTGAAC 758
Db 1720 CTGACCTTTCTCTGACGGAATCTTGGTGGCTACAGCTGAAGATGATGAATTAAGC 1779
QY 759 AAGATTCACAACTGGCTTAAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACCTG 818
Db 1780 CGGACGACCTATTTGGAGCGCATTTCCAGCAGTTTCAAGGACAGAACGATGTATAGG 1839
QY 819 GCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCATGGGCAAACTGTA-----871
Db 1840 GCCTTCAGAGGGAATTTGAACCTAAGAACCTGTAATCATGATCTCTTGAGACTGTA 1899
QY 872 -----TTCACTCAAAACAGATCTTTTCAACACTGAAGATAAGTCAAGTGAACCCAGA 924
Db 1900 CGAATATTTCTGACAGACGAGCCTTTTGAAGGAGACTAGAGAACTCTTACCAGGAGCCCA 1959
QY 925 AGACGGAAGCATGGCTGGA-----TAACTTTGCCCGGT 957
Db 1960 GAGTGCCTCTGAGGAGAGCCCGAGAAATGTCATCGGCTTCTACGAAAGCAGGCTGAG 2019
QY 958 GTTGGGAATATTAGTCCAAAACTTGAA-----AAGAGTACA 995
Db 2020 GAGTCAATATCTGAGTGGGAAAAATTTGAACCTGCATCCGCTGACTGGCAGAGAAAAATA 2079
QY 996 GCACAGACCTTCAAGACTCCAGGAACTTCAAGAGCCACGAGTGGCTGACCTCAAG 1055
Db 2080 GATGAGACCTTGAAGACTCCAGGAACTTCAAGAGCCACGAGTGGCTGACCTCAAG 2139
QY 1056 CTGCGCCAAAGCTGAGGTGATCAAGGGATCTCGGACGCCCGTGGCGGATCTCTCTATTGAC 1115

Db 2140 CTGCGCCAAAGCTGAGGTGATCAAGGGATCTCGGACGCCGTGGCCGATCTCTCATTTGAC 2199
QY 1116 TCTCTCAAGATCACTCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAA 1175
Db 2200 TCTCTCAAGATCACTCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAA 2259
QY 1176 GAGAACCTGAGCCACGTCATGACCTTGTCTGCGCAGCTTACCACTTTGSGCATTCAGCTC 1235
Db 2260 GAGAACCTGAGCCACGTCATGACCTTGTCTGCGCAGCTTACCACTTTGSGCATTCAGCTC 2319
QY 1236 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGTTCCTGAGGTG 1295
Db 2320 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGTTCCTGAGGTG 2379
QY 1296 GCCGTCAGGACCGAGTCAGGCACTGATGAAGCCACACAGGACCTTGGTCCAGCATCT 1355
Db 2380 GCCGTCAGGACCGAGTCAGGCACTGATGAAGCCACACAGGACCTTGGTCCAGCATCT 2439
QY 1356 CAGCACTTTCTTTCCAGCTCTGTCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAA 1415
Db 2440 CAGCACTTTCTTTCCAGCTCTGTCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAA 2499
QY 1416 GTGCCCTACTATATCAACACGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACA 1475
Db 2500 GTGCCCTACTATATCAACACGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACA 2559
QY 1476 GAGCTCTACAGTCTTTTGTAGCTGACCTGAATTAATGTGAGATTTCTCAGCTTATAGACTGCC 1535
Db 2560 GAGCTCTACAGTCTTTTGTAGCTGACCTGAATTAATGTGAGATTTCTCAGCTTATAGACTGCC 2619
QY 1536 ATGAACTCCGAAAGACTGCAAGAGGCCCTTTGTGATCTCTTTGAGCCTGTGAGCTGCA 1595
Db 2620 ATGAACTCCGAAAGACTGCAAGAGGCCCTTTGTGATCTCTTTGAGCCTGTGAGCTGCA 2679
QY 1596 TGTGATGCTTGGACGACGACCAACCTCAAGCAAAATGACCAGCCCATGGATATCTCTGCAG 1655
Db 2680 TGTGATGCTTGGACGACGACCAACCTCAAGCAAAATGACCAGCCCATGGATATCTCTGCAG 2739
QY 1656 ATTATTAATTTGTTGACCACTATTATGACCGCTGAGCAAGAGCAACAAATTTGGTC 1715
Db 2740 ATTATTAATTTGTTGACCACTATTATGACCGCTGAGCAAGAGCAACAAATTTGGTC 2799
QY 1716 AAGTCCCTCTCTGCTGGATATGTCTGAACCTGGCTGCTGAATGTTATGATAGGGA 1775
Db 2800 AAGTCCCTCTCTGCTGGATATGTCTGAACCTGGCTGCTGAATGTTATGATAGGGA 2859
QY 1776 CGAACGAGGAGGATCCGTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCA 1835
Db 2860 CGAACGAGGAGGATCCGTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCA 2919
QY 1836 CATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGATTTGT 1895
Db 2920 CATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGATTTGT 2979
QY 1896 GACCAGCGAGGCTGGGCTCTTCTGATGATTTATCCAAATTCACAGAGTGGGT 1955
Db 2980 GACCAGCGAGGCTGGGCTCTTCTGATGATTTATCCAAATTCACAGAGTGGGT 3039
QY 1956 GAAGTTGCATCTTTTGGGCGAGTAAATTTGAGCCAGTTCGCGA 2001
Db 3040 GAAGTTGCATCTTTTGGGCGAGTAAATTTGAGCCAGTTCGCGA 3085

RESULT 12

US-09-845-416-30
; Sequence 30, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINI-GENE AND USE
; FILE REFERENCE: DB1142

; CURRENT APPLICATION NUMBER: US/09/845,416	
; CURRENT FILING DATE: 2001-04-30	
; PRIOR APPLICATION NUMBER: 60/200,777	
; PRIOR FILING DATE: 2000-04-28	
; NUMBER OF SEQ ID NOS: 36	
; SOFTWARE: PatentIn Ver. 2.1	
; SEQ ID NO 30	
; LENGTH: 4498	
; TYPE: DNA	
; ORGANISM: Homo sapiens	
US-09-845-416-30	
Query Match 62.5%; Score 1251.2; DB 10; Length 4498;	
Best Local Similarity 77.8%; Pred. No. 0;	
Matches 1623; Conservative 0; Mismatches 378; Indels 85; Gaps 6;	
QY	1 GCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTATCAAAACACGCTTTAGAGAA 60
DB	1757 GGCAAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTATCAAAACACGCTTTAGAGAA 1816
QY	61 GTATTATCGTGGCTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 120
DB	1817 GTATTATCGTGGCTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 1876
QY	121 GATGTGGAAGTGTGTAAGACCAAGTTTCACTCATGAGGGGTACATGATGATTTGACA 180
DB	1877 GATGTGGAAGTGTGTAAGACCAAGTTTCACTCATGAGGGGTACATGATGATTTGACA 1936
QY	181 GCCCATCAGGGCGGGTGTGTAATATTCTACAAATGGGAAGTAAGCTGATTTGAAACAGA 240
DB	1937 GCCCATCAGGGCGGGTGTGTAATATTCTACAAATGGGAAGTAAGCTGATTTGAAACAGA 1996
QY	241 AAATTATCAAGATCAAGAACTGAAGTACAGAGCAGATGAATCTCTAAATCAAGA 300
DB	1997 AAATTATCAAGATCAAGAACTGAAGTACAGAGCAGATGAATCTCTAAATCAAGA 2056
QY	301 TGGGAATGCTCAGGCTAGCATGGAAGAAACAAAGCAATTTACATGAGTTTAATG 360
DB	2057 TGGGAATGCTCAGGCTAGCATGGAAGAAACAAAGCAATTTACATGAGTCTATAGA 2116
QY	361 GATCTCCAGAA-----TCGAACTGAAGAGTTCGAATGACTGGTGAACAAACAAAGAA 413
DB	2117 TTAAGTCAACAGTTCCCTGGACCTGGAAAGTTTCTTGGCTGGCTTACAGAGCTGAA 2176
QY	414 GAAAGAACAGGAAATGAGGAGAGCCCTTTGGACCTGATCTTGAAGACCTAAAGCC 473
DB	2177 ACAACTGCCAATGCTCAGGATGCTACCCGTAAGGAAAGGCTCTAGAGACTCCAAG 2236
QY	474 CAAATCAACAACTAAGTGTCTCAAGAGATCTAGAACAAAGCAAGTCAAGGTCAT 533
DB	2237 GGAATAAAGCTGATGAACAAATGGCAAGACCTCCAAGTGAATTTGAAGCTCACACA 2296
QY	534 TCTCTCACTCACATGTTGGTGGTGTGATGA-----ATCTAGTGGAGATCAC 581
DB	2297 GATGTTTATCAACACCTGGATGAAACAGCCAAACAAATCCTGAGATCCCTGGAGGTTCC 2356
QY	582 GCAACTGCTCTTGAAGAACAACTAAG-----TATTGGGAGATCGATGGCAACATC 638
DB	2357 GATGATGCACTGCTGTTCAAGAACGCTTTGGATTAACATGAATTTCAAGTGGAGTGAAT 2416
QY	639 TGTAGATGACAGAACGCTGGGTCTTTTCAAGACATCCTTCTCAATGGCAACGT 698
DB	2417 CGGAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTCTGACCAGTGGAGCGT 2476
QY	699 CTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGAAAGAACAGATGAGTGAAC 758
DB	2477 CTGCACCTTTCTCTGAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGAATTAAGC 2536
QY	759 AAGATTACACAACTGGCTTTTAAAGATCAAAATGAATGTTATCAAGTCTTCAAAACATG 818
DB	2537 CGGAGGCACCTATTGGAGCGACTTTCCAGCAGTTTCAAGACGAGAACGATGTACATAGG 2596
QY	819 GCCGTTTTAAAGCGGATCTAGAAAAAGAAAGCAATTCATGGGCAACTGTA----- 871

DB	2597 GCCTTCAAGAGGGAATTGAANAACCTAAAGAACCTGTATCATGAGTACTCTTGACACTGTA 2656
QY	872 -----TTCACTCAAAACAAGATCTTTTCAACACATGAAGAATAAGTCACTGACCCAGA 924
DB	2657 CGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAACTCTACCAGAGCCACAGA 2716
QY	925 AGACGGAAGCATGCTGGA-----TAACTTTTCCCGGT 957
DB	2717 GAGTGCCTCTGAGGAGAGAGCCAGAAATGTCTACCTGGCTTTTACGAAAGACGAGCTGAG 2776
QY	958 GTTGGGATAATTTAGTCCAAAAAATTGAA-----AAGAGTACA 995
DB	2777 GAGTTCATATCTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 2836
QY	996 GCACAGACCTTTGAAAGACTCCAGAACTTCAAGAGCCAGCGATGAGCTGGACTCAAG 1055
DB	2837 GATGAGACCTTTGAAAGACTCCAGAACTTCAAGAGCCAGCGATGAGCTGGACTCAAG 2896
QY	1056 CTGGCCCAAGCTGAGGTGATCAAGGATCTTGGCAGCCCGTGGCGGATCTCTCATTTGAC 1115
DB	2897 CTGGCCCAAGCTGAGGTGATCAAGGATCTTGGCAGCCCGTGGCGGATCTCTCATTTGAC 2956
QY	1116 TCTCTCAAGATCACTCGAGAAAGTCAAGGACTTTCGAGGAGAAATTTGGCCCTCTGAAA 1175
DB	2957 TCTCTCAAGATCACTCGAGAAAGTCAAGGACTTTCGAGGAGAAATTTGGCCCTCTGAAA 3016
QY	1176 GAGAACTGAGCCACGCTCAATGACTTCTCGCCAGCTTACCACTTTTGGGCAATTCAGCTC 1235
DB	3017 GAGAACTGAGCCACGCTCAATGACTTCTCGCCAGCTTACCACTTTTGGGCAATTCAGCTC 3076
QY	1236 TCACCGTATAACCTCAGACTCTCGAAGACTGTAACACCAAGATGGAAGCTTCTCGAGTG 1295
DB	3077 TCACCGTATAACCTCAGACTCTCGAAGACTGTAACACCAAGATGGAAGCTTCTCGAGTG 3136
QY	1296 GCCCTCGAGGACCGAGTCAGGACCTCATGAAGCCCAAGGACTTTTGGTCCAGCATCT 1355
DB	3137 GCCCTCGAGGACCGAGTCAGGACCTCATGAAGCCCAAGGACTTTTGGTCCAGCATCT 3196
QY	1356 CAGCACTTTCTTCCAGCTCTGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 1415
DB	3197 CAGCACTTTCTTCCAGCTCTGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 3256
QY	1416 GTGCCCTACTATATCAACACGAGACTCAAAACAACTTCTGGGAGCCATCCCAAAATGACA 1475
DB	3257 GTGCCCTACTATATCAACACGAGACTCAAAACAACTTCTGGGAGCCATCCCAAAATGACA 3316
QY	1476 GAGCTCTACCACTCTTTTAGCTGACTGAATAATGTCAGATTTCTCAGCTTATAGACTGCC 1535
DB	3317 GAGCTCTACCACTCTTTTAGCTGACTGAATAATGTCAGATTTCTCAGCTTATAGACTGCC 3376
QY	1536 ATGAAACTCCGAAAGCTTCAGAGAGCCCTTTGCTTGGATCTTTGAGCCTGTCAAGTGCA 1595
DB	3377 ATGAAACTCCGAAAGCTTCAGAGAGCCCTTTGCTTGGATCTTTGAGCCTGTCAAGTGCA 3436
QY	1596 TGTGATGCTTGGACACGACCACTCAAGCAAAATGACCCAGCCCATGGATATCCTCGAG 1655
DB	3437 TGTGATGCTTGGACACGACCACTCAAGCAAAATGACCCAGCCCATGGATATCCTCGAG 3496
QY	1656 ATTATTAATTTGTTTGGACCACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTC 1715
DB	3497 ATTATTAATTTGTTTGGACCACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTC 3556
QY	1716 AACGTCCCTCTCTCGCTGGATATGTCGTAATGCTGGCTGCTGAATGTTTATGATACGGGA 1775
DB	3557 AACGTCCCTCTCTCGCTGGATATGTCGTAATGCTGGCTGCTGAATGTTTATGATACGGGA 3616
QY	1776 CGAACAGGGAGATCCCGTCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 1835
DB	3617 CGAACAGGGAGATCCCGTCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 3676
QY	1836 CATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGATTTTGT 1895

Db 3677 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 3736
Qy 1896 GACCAGCGAGGTGGGCTCCTTCCTGCAAGATTTCTATCCAAATTCGAAGACAGTTGGGT 1955
Db 3737 GACCAGCGAGGTGGGCTCCTTCCTGCAAGATTTCTATCCAAATTCGAAGACAGTTGGGT 3796
Qy 1956 GAAGTTGCATCCTTTGGGGCGAGTACATTGAGCCAGTGTCCGA 2001
Db 3797 GAAGTTGCATCCTTTGGGGCGAGTACATTGAGCCAGTGTCCGA 3842

RESULT 13
US-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-2

Query Match 62.2%; Score 1245; DB 10; Length 4182;
Best Local Similarity 73.1%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 736; Gaps 2;

Qy 1 GGCAGTTTCATGTAGTGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAA 60
Db 1000 GGCAGTTTCATGTAGTGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAA 1059

Qy 61 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTCGAAGCAAGAGAGATTTCTAAT 120
Db 1060 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTCGAAGCAAGAGAGATTTCTAAT 1119

Qy 121 GATGTGAAGTGGTGAAGACCAAGTTTCTACTCATGAGGGTACATGATGGATTGACA 180
Db 1120 GATGTGAAGTGGTGAAGACCAAGTTTCTACTCATGAGGGTACATGATGGATTGACA 1179

Qy 181 GCCCATCAGGCGGGTGGTATATTTCTACAAATGGGAAGTAAAGTGAATGGACACGA 240
Db 1180 GCCCATCAGGCGGGTGGTATATTTCTACAAATGGGAAGTAAAGTGAATGGACACGA 1239

Qy 241 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCCTAAATTCAGA 300
Db 1240 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCCTAAATTCAGA 1299

Qy 301 TGGGAATGCCCTCAGGTAGCTAGCTAGGAAACCAAGCAATTTACATAGATTTTAATG 360
Db 1300 TGGGAATGCCCTCAGGTAGCTAGCTAGGAAACCAAGCAATTTACATAGATTTTAATG 1359

Qy 361 GATCTCCAGAAATC-GAAATCTGAAGAGTTGAATGACTGCTAACAAAAACAGAAAGA 419
Db 1360 GATCTCCAGAAATCAGAAATCTGAAGAGTTGAATGACTGCTAACAAAAACAGAAAGA 1419

Qy 420 ACAAGGAAATGAGGAAGAGCTCTTGGACCTGATCTTTGAGACCTTAAACGCCAAGTA 479
Db 1420 ACAAGGAAATGAGGAAGAGCTCTTGGACCTGATCTTTGAGACCTTAAACGCCAAGTA 1479

Qy 480 CAACACATAGGTGCTTCAAGAGAGATCTAGAACAGAAACAAGTCAAGGTCAATTTCTCTC 539
Db 1480 CAACACATAGGTGCTTCAAGAGAGATCTAGAACAGAAACAAGTCAAGGTCAATTTCTCTC 1539

Qy 540 ACTCACATGTTGGTGGTGTGATGAATCTAGTGGAGATCACGCCAACTGCTGTTGGAA 599

Db 1540 ACTCACATGTTGGTGGTGTGATGAATCTAGTGGAGATCACGCCAACTGCTGCTTTGGAA 1599
Qy 600 GAACAACTTAAGGTATTGGAGATTCGATGGGCAAAACATCTGTAGATGGACAGAACCCG 659
Db 1600 GAACAACTTAAGGTATTGGAGATTCGATGGGCAAAACATCTGTAGATGGACAGAACCCG 1659
Qy 660 TGGTCTCTTTTCAAGACATCCTTCTCAATGCAAGCTCTTACTGAAAGCAAGTGCCTT 719
Db 1660 TGGTCTCTTTTCAAGACATCCTTCTCAATGCAAGCTCTTACTGAAAGCAAGTGCCTT 1719
Qy 720 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCAAGTGAACAAAGATTCAACAACTGGCTT 779
Db 1720 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCAAGTGAACAAAGATTCAACAACTGGCTT 1779
Qy 780 AAGATCAAAATGAATGTTTATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTA 839
Db 1780 AAGATCAAAATGAATGTTTATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTA 1839
Qy 840 GAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTCACTCAAAACAGATCTTCTTTCAACA 899
Db 1840 GAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTCACTCAAAACAGATCTTCTTTCAACA 1899
Qy 900 CTGAAGAAATAGTCAAGTACCCAGAGACGGAAGCATGGCTGATTAACCTTGGCCCGTGT 959
Db 1900 CTGAAGAAATAGTCAAGTACCCAGAGACGGAAGCATGGCTGATTAACCTTGGCCCGTGT 1959
Qy 960 TGGGATAATTTAGTCCAAAAAATTTGAAAAGAGTACAGCAC- 999
Db 1960 TGGGATAATTTAGTCCAAAAAATTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAA 2019
Qy 1000 ----- 999
Db 2020 CAGTTCCCTCGACCTGGAAAAAGTTTCTTCTGCTTACAGAACTGAAACAACTGCC 2079
Qy 1000 ----- 999
Db 2080 AATGCTCTACAGGATGCTACCCGTAAGGAAAGCTCCTAGAAAGCTTCCAAAGGAGTAA 2139
Qy 1000 ----- 999
Db 2140 GAGCTGATGAACAAATGGCAAGACCTTCAAGGTGAAATGAAGCTCACACAGATGTTAT 2199
Qy 1000 ----- 999
Db 2200 CACAACCTGGATGAAAAACAGCCAAAAAATCTGAGATCCCTGGAAAGTTCGAGATGCA 2259
Qy 1000 ----- 999
Db 2260 GTCTGTTCAAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGNACTTCGAAAAAG 2319
Qy 1000 ----- 999
Db 2320 TCTCTCAACATTAGGTCCCATTTTGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTT 2379
Qy 1000 ----- 999
Db 2380 TCTCTGAGGAACCTTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGCA 2439
Qy 1000 ----- 999
Db 2440 CCTATTGGAGCGACTTTCCAGCAGTTTCAAGAGCAAGATGATACATAGGGCCTTCAAG 2499
Qy 1000 ----- 999
Db 2500 AGGAAATTGAAAATAAAGAACTGTATCATGAGTACTCTTGAGACTGTACGAATATTT 2559
Qy 1000 ----- 999
Db 2560 CTGACAGAGCAGCCTTTTGAAGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCT 2619
Qy 1000 ----- 999

2417	Db	TGGGTTCTTTTACAAGACATCCTTCTCAAAATGGCAACGCTTTACTAGTAAGAACAGTGCCTT	2476
720	Qy	TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCACTGAAACAAGATTCAACAACCTGGCTTT	779
2477	Db	TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCACTGAAACAAGATTCAACAACCTGGCTTT	2536
780	Qy	AAAGATCAAAATGAATGTTATCAAGTCTTCAAAACTGGCCGTTTTTAAAGCGGATCTA	839
2537	Db	AAAGATCAAAATGAATGTTATCAAGTCTTCAAAACTGGCCGTTTTTAAAGCGGATCTA	2596
840	Qy	GAAGAAGAAAGCAATCCATGGGCAAACTGTATTCCTCAAAACAAGATCTTCTTCAACA	899
2597	Db	GAAGAAGAAAGCAATCCATGGGCAAACTGTATTCCTCAAAACAAGATCTTCTTCAACA	2656
900	Qy	CTGAAGAATAAGTCAGTGACCCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCCGTTG	959
2657	Db	CTGAAGAATAAGTCAGTGACCCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCCGTTG	2716
960	Qy	TGGGTAATTTAGTCCAAAACCTTGAAGAAGTACAGCAC-----	999
2717	Db	TGGGTAATTTAGTCCAAAACCTTGAAGAAGTACAGCAC-----	2776
1000	Qy	-----	999
2777	Db	CAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAGCTGAACAACCTGCC	2836
1000	Qy	-----	999
2837	Db	AATGTCTCAGGATGCTACCCGTAAGAAAAAGGCTCTAGAGACTTCMAGGGAGTAAAA	2896
1000	Qy	-----	999
2897	Db	GAGCTGATGAAACAATGGCAAGACCTCCAGGTGAAATTCGAAGCTCACAGATGTTAT	2956
1000	Qy	-----	999
2957	Db	CACAACTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAAGTTCGGATGATGCA	3016
1000	Qy	-----	999
3017	Db	GTCTGTTCAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAATTCGGAAAAAG	3076
1000	Qy	-----	999
3077	Db	TCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTT	3136
1000	Qy	-----	999
3137	Db	TCTCTGCAGGAACTTCTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCA	3196
1000	Qy	-----	999
3197	Db	CCTATTGGAGGGCAGCTTCCAGCAGTTTCAGAACGATCTACATAGGGCCTTCAAG	3256
1000	Qy	-----	999
3257	Db	AGGGAAATTGAAAACCTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTT	3316
1000	Qy	-----	999
3317	Db	CTGACAGACAGCCTTTGGAAAGCACTAGAGAAACTCTACAGGAGCCCAGAGAGTGCCT	3376
1000	Qy	-----	999
3377	Db	CCTGAGGAGAGAGCCAGAATGTCACTCGGCTTCTACGAAAGCAGCGCTGAGGAGTCAAT	3436
1000	Qy	-----AGACC	1004
3437	Db	ACTGAGTGGAAAAATTGAACCTTGCACTCGCTGACTGGCAGAGAAAAATAGATGAGACC	3496
1005	Qy	CTTGAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAA	1064
3497	Db	CTTGAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAA	3556

Qy	1065	GCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGGGATCTCCTCAATTGACTCTCTCCAA	1124
Db	3557	GCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGGGATCTCCTCAATTGACTCTCTCCAA	3616
Qy	1125	GATCACCTCGAGAAAGTCAAGGCACATTGAGGAGAAATTGCGCTCTGAAAGAGAAGCTG	1184
Db	3617	GATCACCTCGAGAAAGTCAAGGCACATTGAGGAGAAATTGCGCTCTGAAAGAGAAGCTG	3676
Qy	1185	AGCCACGTCAATGACCTTGTCTGCCAGCTTACCACTTTTGGGCAATCAGCTCTCACCGTAT	1244
Db	3677	AGCCACGTCAATGACCTTGTCTGCCAGCTTACCACTTTTGGGCAATCAGCTCTCACCGTAT	3736
Qy	1245	AACCTCAGCACTCTGGAAGACCTGAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAG	1304
Db	3737	AACCTCAGCACTCTGGAAGACCTGAACACAGATGGAAGCTTCTGCAGGTGGCCGTCGAG	3796
Qy	1305	GACCGAGTCAGGAGCTGATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGCACTTT	1364
Db	3797	GACCGAGTCAGGAGCTGATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGCACTTT	3856
Qy	1365	CTTTCCACGTCCTGCAGGCTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCTTAC	1424
Db	3857	CTTTCCACGTCCTGCAGGCTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCTTAC	3916
Qy	1425	TATATCAACCCAGAGACTCAAAACAATTGCTGGGACCATCCCAAAATGACAGAGCTCTAC	1484
Db	3917	TATATCAACCCAGAGACTCAAAACAATTGCTGGGACCATCCCAAAATGACAGAGCTCTAC	3976
Qy	1485	CAGCTTTAGCTGACCTGAATATATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTC	1544
Db	3977	CAGCTTTAGCTGACCTGAATATATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTC	4036
Qy	1545	CGAAGACTGCAGAAGGCCCTTTGCTTTGGATCTCTTGAGCCTGTGACGTGATGATGCC	1604
Db	4037	CGAAGACTGCAGAAGGCCCTTTGCTTTGGATCTCTTGAGCCTGTGACGTGATGATGCC	4096
Qy	1605	TTGGACCGACCAAACTCAAGCAAAATGACGCCCATGGATATCTCGCAGATATTATAT	1664
Db	4097	TTGGACCGACCAAACTCAAGCAAAATGACGCCCATGGATATCTCGCAGATATTATAT	4156
Qy	1665	TGTTTCACCACTATTATGACCCGCTGGAGCAGACCAAAATTTGCTCAACGTCCT	1724
Db	4157	TGTTTCACCACTATTATGACCCGCTGGAGCAGACCAAAATTTGCTCAACGTCCT	4216
Qy	1725	CTCTGGGTGGATATGTCGTAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGG	1784
Db	4217	CTCTGGGTGGATATGTCGTAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGG	4276
Qy	1785	AGGATCCGTGCTCTGCTTTTAAACTGGCATTCTCCCTGTGTAAGACAAATTTGGAA	1844
Db	4277	AGGATCCGTGCTCTGCTTTTAAACTGGCATTCTCCCTGTGTAAGACAAATTTGGAA	4336
Qy	1845	GACAAGTACAGATACCTTTTAAAGCAAGTGGCAAGTTCAAACAGGATTTTGTGACACGCG	1904
Db	4337	GACAAGTACAGATACCTTTTAAAGCAAGTGGCAAGTTCAAACAGGATTTTGTGACACGCG	4396
Qy	1905	AGGCTGGGCTCCTCTGTCATGATTCATCCAAATTCCAAGACAGTTGGGTGAAGTTGCA	1964
Db	4397	AGGCTGGGCTCCTCTGTCATGATTCATCCAAATTCCAAGACAGTTGGGTGAAGTTGCA	4456
Qy	1965	TCCTTTGGGGCAGTAACATTTGAGCCAAAGTGC CGGA	2001
Db	4457	TCCTTTGGGGCAGTAACATTTGAGCCAAAGTGC CGGA	4493

```

RESULT 15
US-10-149-736-41
; Sequence 41, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott O.

```

; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences			
; FILE REFERENCE: UM-06968			
; CURRENT APPLICATION NUMBER: US/10/149,736			
; CURRENT FILING DATE: 2002-06-17			
; PRIOR APPLICATION NUMBER: PCT/US01/31126			
; PRIOR FILING DATE: 2001-10-04			
; PRIOR APPLICATION NUMBER: 60/238,848			
; PRIOR FILING DATE: 2000-10-06			
; NUMBER OF SEQ ID NOS: 96			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 41			
; LENGTH: 5462			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Synthetic			
US-10-149-736-41			
Query Match 56.6%; Score 1132.2; DB 16; Length 5462;			
Best Local Similarity 73.9%; Pred. No. 0;			
Matches 1628; Conservative 0; Mismatches 373; Indels 202; Gaps 7;			
Qy	1	GGCAGTTCAATGATGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAA	60
Db	1199	GGCAGTTCAATGATGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAA	1258
Qy	61	GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAGGAGAGATTTCTAAT	120
Db	1259	GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAGGAGAGATTTCTAAT	1318
Qy	121	GATGTGGAAGTGTGAAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTGACA	180
Db	1319	GATGTGGAAGTGTGAAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTGACA	1378
Qy	181	GCCCATCAGGCGCGGTGGTATATTTCTACATTTGGGAAGTAAGCTGATGGACAGGA	240
Db	1379	GCCCATCAGGCGCGGTGGTATATTTCTACATTTGGGAAGTAAGCTGATGGACAGGA	1438
Qy	241	AAATTATCAGAAGATGAAGAAATCGAAGTACAGAGCAGATGAATCTCCTTAATTCAAGA	300
Db	1439	AAATTATCAGAAGATGAAGAAATCGAAGTACAGAGCAGATGAATCTCCTTAATTCAAGA	1498
Qy	301	TGGGAATGCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGATTTTAATG	360
Db	1499	TGGGAATGCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATGCTCTGGACTG	1558
Qy	361	-----	360
Db	1559	ACCACATTTGGAGCCTCTCCTACTCAGACTGTTACTCTGCTGACACAACTGTGGTTACT	1618
Qy	361	-----GATCTCCAGAAATCGAAA-----	377
Db	1619	AAGGAACTGCCATCTCCAAACTAGAAAATGCCATCTTCTTGATGTTGGAGCATAGATTA	1678
Qy	378	-----CTGAAGAGTTGAATGACTGGCTAACAACCAAGACAGAGAA	416
Db	1679	CTGCAACAGTTCCCTCGTGGACCTGGAAAAAGTTTCTTGGCTGGTTACAGAAGCTGAACA	1738
Qy	417	AGAACAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAA	476
Db	1739	ACTGCCAATGCTCTACAGGATGTACCCGTAAAGAAAGGCTCTAGAGACTCCAAAGGA	1798
Qy	477	GTACACAAACATAGGTGCTTCAAGAGATCTAGAACAGAAACAGTCAAGGTCAATTC	536
Db	1799	GTAAAAGAGTGTGAACAAATGGCAAGCCTCCAAGGTGAATTTGAAGCTCACACAGAT	1858
Qy	537	CTCACTCACATGTGGTGGTGGTATGATGA-----ATCTAGTGGAGATCACGCA	584
Db	1859	GTTTATCACAACTGGATGAACAGCCAAATAATCCTGNGATCCCTGGAAAGTTCCGAT	1918
Qy	585	ACTGCTGCTTTGGAAGAACAACTTAAGG---TATTGGAGATCGATGGGCAACATCTCT	641
Db	1919	GATGCACTGCTGTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAATCTCGG	1978

Qy	642	AGATGGACAGAAAGACCCCTGGTTCTTTTACAGACATCCTTCTCAATGGCAAGTCTTT	701
Db	1979	AAAAAGTCTCTCAACATTAGTCTCCCATTTGGAAGCCAGTTCTTGACCAAGTGAACGCTG	2038
Qy	702	ACTGAAGAAACAGTGCCTTTTGTGTCATGGCTTTTTCAGAAAAAGAAAGTGAAGTGAAG	761
Db	2039	CACCTTTCTCTGCAGGAACCTTCTGCTGCTCAGCTGAAGATGATGAATTAAGCCGG	2098
Qy	762	ATTCAACAACCTGCTTTTAAAGATCAAAATGAATGTTTCAAGTCTTTCAAAAACCTGGCC	821
Db	2099	CAGGCACCTATTGGAGCGACTTTTCCAGCAGTTTTCAGAAAGCAGACGATGTACATAGGCC	2158
Qy	822	GTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCATGGGCAAACTGTA-----	871
Db	2159	TTCAAGAGGGAATTGAAAAAATTAAGAAACCTGTATATCATGAGTACTCTTTGAGACTGTAGA	2218
Qy	872	---TTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAAGTCAAGCCAGAGA	927
Db	2219	ATATTTCTGACAGAGAGCCCTTTGGAAGGACTAGAGAAACTCTACCAAGGACCCAGAGAG	2278
Qy	928	CGGAAGCATGCTCGA-----TAACTTTGCCCGGTGTT	960
Db	2279	CTGCTCTGAGGAGAGAGCCAGAAATGTCACTCGGTTTACGAAAAGCAGGCTGAGAGAG	2338
Qy	961	GGGATAATTAGTCCAAAAAATTTGAA-----AAGAGTACAGCA	998
Db	2339	GTCAATCTGATGGGAAAAATTTGAACCTGCATCCGCTGACTGGCAGAGAAAAATAGAT	2398
Qy	999	CAGACCTTTGAAAGACTTCAGGAACTTCAAGAGGCCACGATGAGCTGGACCTCAAGCTG	1058
Db	2399	CAGACCTTTGAAAGACTTCAGGAACTTCAAGAGGCCACGATGAGCTGGACCTCAAGCTG	2458
Qy	1059	CGCAAGCTGAGGTGATCAAGGGATCTGCGAGCCCTGGCGCATCTCTCATTTGACTCT	1118
Db	2459	CGCAAGCTGAGGTGATCAAGGGATCTGCGAGCCCTGGCGCATCTCTCATTTGACTCT	2518
Qy	1119	CTCAAGATCACTCGAAGAAAGTCAAGCACTTCGAGGAGAAATTTGCGCTCTCGAAAGAG	1178
Db	2519	CTCAAGATCACTCGAAGAAAGTCAAGCACTTCGAGGAGAAATTTGCGCTCTCGAAAGAG	2578
Qy	1179	AAGCTGACCACTCAATGACCTTCTGCGCAGCTTACCACTTTGGGCATTCAGCTCTCA	1238
Db	2579	AAGCTGACCACTCAATGACCTTCTGCGCAGCTTACCACTTTGGGCATTCAGCTCTCA	2638
Qy	1239	CCGTATAACCTCAGCACTCTGGAAGAGCTGAACACCAAGATGGAAAGCTTTCTGAGGTGGCC	1298
Db	2639	CCGTATAACCTCAGCACTCTGGAAGAGCTGAACACCAAGATGGAAAGCTTTCTGAGGTGGCC	2698
Qy	1299	GTGAGGACCGAGTCCAGGACCTGCATGAAGCCCAAGAGGACCTTTGGTCCAGCATCTCAG	1358
Db	2699	GTGAGGACCGAGTCCAGGACCTGCATGAAGCCCAAGAGGACCTTTGGTCCAGCATCTCAG	2758
Qy	1359	CACCTTTCTTCCAGCTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTG	1418
Db	2759	CACCTTTCTTCCAGCTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTG	2818
Qy	1419	CCCTACTATATCAACCAAGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATCAAGAG	1478
Db	2819	CCCTACTATATCAACCAAGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATCAAGAG	2878
Qy	1479	CTCTACAGTCTTTAGTCTGACCTGAATAATGTGAGATTTCTCAGCTTATAGGACTGCCATG	1538
Db	2879	CTCTACAGTCTTTAGTCTGACCTGAATAATGTGAGATTTCTCAGCTTATAGGACTGCCATG	2938
Qy	1539	AAACTCCGAAGACTGCAAGAGGCCCTTTGCTTGGATCTCTTGGCCCTGTGAGCTGCATGT	1598
Db	2939	AAACTCCGAAGACTGCAAGAGGCCCTTTGCTTGGATCTCTTGGCCCTGTGAGCTGCATGT	2998
Qy	1599	GATGCTTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGGATATCTTCAGATT	1658
Db	2999	GATGCTTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGGATATCTTCAGATT	3058

```
Qy 1659 ATTAATTGTTTGACCACTATTATGACCGCTGGAGCAAGACACAACAAATTTGGTCAAC 1718
Db |||||||
Qy 3059 ATTAATTGTTTGACCACTATTATGACCGCTGGAGCAAGACACAACAAATTTGGTCAAC 3118
Db |||||||
Qy 1719 GTCCCTCTCTGCGTGGATATGCTCTGAACTGGCTGCTGAAATGTTTATGATACGGGACGA 1778
Db |||||||
Qy 3119 GTCCCTCTCTGCGTGGATATGCTCTGAACTGGCTGCTGAAATGTTTATGATACGGGACGA 3178
Db |||||||
Qy 1779 ACAGGGAGGATCCGTGTCCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAGCACAT 1838
Db |||||||
Qy 3179 ACAGGGAGGATCCGTGTCCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAGCACAT 3238
Db |||||||
Qy 1839 TTGGAAGACAAAGTACAGATACCTTTTCAAGCAAAGTGGCAAGTTCAAACAGGATTTTGTGAC 1898
Db |||||||
Qy 3239 TTGGAAGACAAAGTACAGATACCTTTTCAAGCAAAGTGGCAAGTTCAAACAGGATTTTGTGAC 3298
Db |||||||
Qy 1899 CAGCGCAGGCTGGGCGCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAA 1958
Db |||||||
Qy 3299 CAGCGCAGGCTGGGCGCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAA 3358
Db |||||||
Qy 1959 GTTGCAATCCTTTGGGGGCGAGTAAACATTGAGCCAAAGTGTCCGGA 2001
Db |||||||
Qy 3359 GTTGCAATCCTTTGGGGGCGAGTAAACATTGAGCCAAAGTGTCCGGA 3401
Db |||||||
```

Search completed: September 19, 2004, 15:15:02
Job time : 710.733 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 19, 2004, 02:48:31 ; Search time 3619.24 Seconds
(without alignments)
16510.166 Million cell updates/sec

Title: US-09-845-416-14_COPY_1000_3000

Perfect score: 2001

Sequence: 1 ggcagtcattgatgagag.....cattgagccaaagtgtccga 2001

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

5 571.4 28.6 5697 29 AY399455
6 569 28.4 777 14 CD653550
7 569 28.4 824 9 AL556247
8 567.4 28.4 1098 13 BX365572
9 566.4 28.3 801 14 CB991394
10 558 27.9 620 13 BQ640063
11 532.2 28.6 770 12 BG719710
12 511.4 25.6 1298 11 AK087829
13 511.4 25.6 2135 11 AK013510
14 511.4 25.6 4437 11 AK036936
15 511 25.5 797 14 CB960722
16 509.8 25.5 1384 11 AK075809
17 502 25.1 5676 29 AY399454
18 494.6 24.7 652 10 BB629984
19 491.6 24.6 704 10 BB610411
20 489.8 24.5 508 14 CD701871
21 481.2 24.0 1047 14 CB850319
22 467.2 23.3 854 9 A1196693
23 460.6 23.0 3753 11 AK081426
24 457.2 22.8 513 29 CG605968
25 455.8 22.8 591 13 BX490860
26 455.8 22.8 717 14 CB527785
27 453.8 22.7 508 29 CG606947
28 446.2 22.3 542 29 CG594848
29 439 21.9 578 13 BX485574
30 436.6 21.8 665 13 BY742604
31 427.4 21.4 2874 29 AY408546
32 422.8 21.1 898 10 BF182065
33 422.6 21.1 520 29 CG606443
34 422.6 21.1 2874 29 AY408548
35 420.2 21.0 512 14 CF162938
36 419.4 21.0 488 29 CG597965
37 417.6 20.9 843 14 CA988247
38 416 20.8 493 29 CG604818
39 410.4 20.5 599 10 BB666688
40 407.2 20.3 554 14 CB613696
41 405.4 20.3 662 12 BJ626491
42 404.4 20.2 493 14 CA888041
43 397.4 19.9 495 14 CA894775
44 397 19.8 595 14 CB177816
45 392.2 19.6 650 13 BY174491

ALIGNMENTS

RESULT 1
BC036103
LOCUS
DEFINITION
Homo sapiens dystrophin (muscular dystrophy, Duchenne and Becker types), mRNA (cdna clone IMAGE:5274415), with apparent retained intron.
3870 bp
mRNA
linear
HTC 19-NOV-2003

ACCESSION
BC036103
VERSION
BC036103.1
KEYWORDS
HTC.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens (human)
REFERENCE
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 3870)
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.P., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carrinici,P., Prange,C., Raha,S.S., Joqueillano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahsey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	991	49.5	3870	11 BC036103	BC036103 Homo sapi
2	776.6	38.8	3056	11 AK044536	AK044536 Mus muscu
3	765	38.2	5691	29 AY399453	AY399453 Homo sapi
4	585.2	29.2	728	14 CB228986	CB228986 AGENCOURT

Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M., Butterfield, Y. S., Krzywicki, M. I., Skalska, U., Smalish, D. E., Schnerch, A., Schein, J. E., Jones, S. J. and Marra, M. A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2388257
12477932
2 (bases 1 to 3870)
Strausberg, R.
Direct Submission
Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: ccaphs@email.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) medpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 48 Row: f Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
1..3870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5274415"
/tissue_types="Brain, hippocampus"
/clone_lib="NIH MGC_95"
/lab_host="DH10B"
/note="vector: pBluescript"

ORIGIN

Query Match 49.5%; Score 991; DB 11; Length 3870;
Best Local Similarity 99.9%; Pred. No. 6.6e-223;
Matches 1002; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 1 GGCAGTTCATGTGGAGAGTGAAGTAACTGGACCGTTATCAACACAGCTTTAGAGAA 60
Db 1143 GGCAGTTCATGTGGAGAGTGAAGTAACTGGACCGTTATCAACACAGCTTTAGAGAA 1202
Qy 61 GTATTATCGTGGCTTCCTTTCTGTGGAGACACATTCGACACACAGGAGAGATTCTAAT 120
Db 1203 GTATTATCGTGGCTTCCTTTCTGTGGAGACACATTCGACACACAGGAGAGATTCTAAT 1262
Qy 121 GATGTGAAGTGTGAAGACACAGCTTCATCATCATGAGGGTACATGATGATTGACA 180
Db 1263 GATGTGAAGTGTGAAGACACAGCTTCATCATCATGAGGGTACATGATGATTGACA 1322
Qy 181 GCCCATCAGGCGCGGTGGTGTATATTTCTACATTTGGAACTAAGCTGATTGGAACAGGA 240
Db 1323 GCCCATCAGGCGCGGTGGTGTATATTTCTACATTTGGAACTAAGCTGATTGGAACAGGA 1382
Qy 241 AATTATCAGAAGATGAAGAACTGAAGTCAAGAGCAGATGAATCTCTTAATTCAGA 300
Db 1383 AATTATCAGAAGATGAAGAACTGAAGTCAAGAGCAGATGAATCTCTTAATTCAGA 1442

Qy 301 TGGGAATGCTCAGGGTAGCTAGCATGGAATAAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 1443 TGGGAATGCTCAGGGTAGCTAGCATGGAATAAAACAAAGCAATTTACATAGAGTTTAAATG 1502
Qy 361 GATCTCCAGATC-GAAACTGAAAGAGTGAATGACCTGGCTTAACAAAAACAGAGAAAGA 419
Db 1503 GATCTCCAGATCAGAAACTGAAAGAGTGAATGACCTGGCTTAACAAAAACAGAGAAAGA 1562
Qy 420 ACAAGGAAATGGAGGAGAGCCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTA 479
Db 1563 ACAAGGAAATGGAGGAGAGCCCTCTTGGACCTGATCTTGAAGACCTAAACCGCAAGTA 1622
Qy 480 CAACAACATAGGTGCTTCAAGAGATCTAGAACAAGCAAGTCAAGGTCMAATTCCTC 539
Db 1623 CAACAACATAGGTGCTTCAAGAGATCTAGAACAAGTCAAGGTCMAATTCCTC 1682
Qy 540 ACTCACATGCTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCAACTGTCTTTGGA 599
Db 1683 ACTCACATGCTGGTGGTAGTTGATGAATCTAGTGGAGATCAGCAACTGTCTTTGGA 1742
Qy 600 GAACAACTTAAGTATTGGGAGATCGATGGCAACAACATCTGTAGATGGACAGAGACCG 659
Db 1743 GAACAACTTAAGTATTGGGAGATCGATGGCAACAACATCTGTAGATGGACAGAGACCG 1802
Qy 660 TGGGTCTTTTACAAGACATCTTCTCAATGGCAAGCTCTTACTGAAGAACAGTGCCT 719
Db 1803 TGGGTCTTTTACAAGACATCTTCTCAATGGCAAGCTCTTACTGAAGAACAGTGCCT 1862
Qy 720 TTTAGTGCATCGCTTTCAGAAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 779
Db 1863 TTTAGTGCATCGCTTTCAGAAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1922
Qy 780 AAAGTCAAAATGAATGTTATCAAGTCTTCAAAAGTGGCGGTTTTAAAGCGGATCTA 839
Db 1923 AAAGTCAAAATGAATGTTATCAAGTCTTCAAAAGTGGCGGTTTTAAAGCGGATCTA 1982
Qy 840 GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTACTCAAAAGATCTCTTTTCAACA 899
Db 1983 GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTACTCAAAAGATCTCTTTTCAACA 2042
Qy 900 CTGAAGATAAGTCAAGTGAACCCAGAACGAGCATGGCTGATTAATCTTGGCGGTGT 959
Db 2043 CTGAAGATAAGTCAAGTGAACCCAGAACGAGCATGGCTGATTAATCTTGGCGGTGT 2102
Qy 960 TGGGATTAATTTAGTCCAAAACTTGAAGAGAGTACAGACAGA 1002
Db 2103 TGGGATTAATTTAGTCCAAAACTTGAAGAGAGTACAGACAGA 2145

RESULT 2
AK044536
LOCUS
DEFINITION
Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone: A930019F21 product: dystrophin, muscular dystrophy, full insert sequence.
ACCESSION
AK044536
VERSION
AK044536.1
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
REFERENCE
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20493374
11042159

JOURNAL
MEDLINE
PUBMEDREFERENCE
AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
Riken integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

TITLE

JOURNAL
MEDLINE
PUBMEDREFERENCE
AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
11076861

JOURNAL
MEDLINE
PUBMEDREFERENCE
AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
11076861

JOURNAL
MEDLINE
PUBMEDREFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, T., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaki-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission

JOURNAL
MEDLINE
PUBMEDREFERENCE
AUTHORS

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Retina RNA was provided by Dr. Stefano Gustincich (Department of
Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
02115, USA) whose assistance is gratefully acknowledged. Please
visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/
Location/Qualifiers

FEATURES

source

1. .3056
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:A930019F21"
/db_xref="MGI:2409854"
/db_xref="taxon:10090"
/clone="A930019F21"
/tissue_type="retina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1. .3056

misc_feature

/note="dyatrophin, muscular dystrophy (MGI:94909,
GB|NM_007868, evidence: BLASTN, 100%, match=99%)"

ORIGIN

Query Match 38.8%; Score 776.6; DB 11; Length 3056;
Best Local Similarity 86.9%; Pred. No. 2.7e-172; Indels 1; Gaps 1;
Matches 868; Conservative 0; Mismatches 134;
QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTTAAAGAA 60
Db 1044 GACAGTTTCATTGATGGAGAGCGGAAGTAAATCTGGATAGTTACCAAACTGCTTTAAGAA 1103
QY 61 GTATATTCGTCGCTTCCTTCCTGCTGAGACACATTCGACGACAGGAGAGATTTCTAAT 120
Db 1104 GTACTTTTCATGGCTTCCTTCCTGCGGAGATACATTCGAGACCAAGGAGAGATTTCAAT 1163
QY 121 CATGTGGAAGTGTGAAAGACCAAGTTCATCTCATGAGGGGTACATGATGGATTTGACA 180
Db 1164 CATGTTGGAAGTGTGAAAGACCAAGTTCATCTCATGAGGGATTCATGATGGATTTGACA 1223
QY 181 GCCCATCAGGCGCGGTTGGTAATTTCTCAATTTGGAGAGTAAAGCTGATTTGGAACAGA 240
Db 1224 TCTCATCAAGGACTTGTGGTAAATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAGAGG 1283
QY 241 AAATTCAGAGAGATGAAGAACTGAAGTCAAGACGAGATGAATCTCTTAAATTTCAAGA 300
Db 1284 AAATTCAGAGAGATGAAGAACTGAAGTCAAGACCAAGTGAATCTCTTAAATTTCAAGA 1343
QY 301 TGGGAATTCCTCAGGCTAGCTAGCATGAGAAACCAAGCAATTTACATAGAGTTTAAATG 360
Db 1344 TGGGAATTCCTCAGGCTAGCTAGCATGAGAAACCAAGCAATTTACATAGAGTTTAAATG 1403
QY 361 GATCTCCAGATC-GAACTCAAGAGTTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 419
Db 1404 GATCTCCAGATC-GAACTCAAGAGTTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1463
QY 420 ACAAGGAAATTCGAGAGAGAGCTCTTGGAGCTGATCTTGAAGACCTTAAACCGCAAGTA 479
Db 1464 ACTAAGAAATTCGAGAGAGAGCTCTTGGAGCTGATCTTGAAGACCTTAAACCGCAAGTA 1523
QY 480 CAACCAATAAGGTGCTTCAAGAGATCTAGAACAGAACAGTCAAGTCAAGTCAAGTCAAGT 539
Db 1524 CAACCAATAAGGTGCTTCAAGAGATCTAGAACAGAACAGTCAAGTCAAGTCAAGTCAAG 1583
QY 540 ACTCATAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 599
Db 1584 ACTCATAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1643
QY 600 GAACAACTTAAGGTATTCGGAGATCGATGGCAACCAATCTCTAGATGACAGAGAGACCG 659
Db 1644 GAACAACTTAAGGTATTCGGAGATCGATGGCAACCAATCTCTAGATGACAGAGACCG 1703
QY 660 TGGGTTCTTTTACAGACATCTCTCAATGGCAAGCTCTTACTGAGAGACAGTGCCTT 719
Db 1704 TGGATGTTTTCAGAGATATTTCTTAAATGGCAGCATTTTCTGAGAGACAGTGCCTT 1763
QY 720 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCAGTGAACAGAGATTCACAACTGGCTTT 779
Db 1764 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCAGTGAACAGAGATTCACAACTGGCTTT 1823
QY 780 AAAGATCAAAATGAATGTTATCAAGTCTTCAAAACCTGGCGCTTTTAAAGCGGATCTA 839
Db 1824 AAAGATCAAAATGAATGTTATCAAGTCTTCAAAACCTGGCGCTTTTAAAGCGGATCTA 1883
QY 840 GAAAGAAAAAGCAATTCATGGCAAACTGTTTCACTCAAAACAGATCTTCTTTCAACA 899
Db 1884 GAAAGAAAAAGCAATTCATGGCAAACTGTTTCACTCAAAACAGATCTTCTTTCAACA 1943
QY 900 CTGAGAAATTAAGTCAAGTCAAGCAAGAGACGGAAGATCGGTGGATTAATTTGCGCGGTGT 959
Db 1944 CTGAGAAATTAAGTCAAGTCAAGCAAGAGATCGGTGGATTAATTTGCGCGGTGT 2003
QY 960 TGGGATAATTTAGTCCAAAAAATTTGAAAGAGTACAGCACAGA 1002

```

Db 2004 TGGGACAAATTAAACCCAAAACCTGAAAGAGTTGACGACAAA 2046
|||||
RESULT_3
AY399453
LOCUS
DEFINITION Homo sapiens HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
AY399453 genomic survey sequence.
ACCESSION
VERSION AY399453.1 GI:39755442
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 5691)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
2 (bases 1 to 5691)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..5691
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..5691
/locus_tag="HCM0229"
gene
ORIGIN
Query Match 38.2%; Score 765; DB 29; Length 5691;
Best Local Similarity 89.1%; Pred. No. 1.7e-169;
Matches 894; Conservative 0; Mismatches 0; Indels 109; Gaps 2;
QY 1 GGCAGTTTCATGTGAGAGTGAAGTAAACCTGGACCGTTATCAACACGCTTTAGAGAA 60
Db 385 GGCAGTTTCATGTGAGAGTGAAGTAAACCTGGACCGTTATCAACACGCTTTAGAGAA 444
QY 61 GTATTATCGTGGCTTTCTTCTGCTGAGGACACATTCGACGACCAAGGAGAGATTCTTAAT 120
Db 445 GTATTATCGTGGCTTTCTTCTGCTGAGGACACATTCGACGACCAAGGAGAGATTCTTAAT 504
QY 121 GATGTGGAATGTGTAAGAACACCACTTCTATCTATCTAGGGGTATCATGATGATTGACA 180
Db 505 GATGTGGAATGTGTAAGAACCACTTCTATCTATCTAGGGGTATCATGATGATTGACA 564
QY 181 GCCCATCAGGCGCGTGGTATATTTCTACAAATGGGAAGTAAAGTATTCGACGACGAGA 240
Db 565 GCCCATCAGGCGCGTGGTATATTTCTACAAATGGGAAGTAAAGTATTCGACGACGAGA 624
QY 241 AAATTATCAGAGATGAGAAACTGAAGTACAGGACGACATGAATCTCTTAATTCAGAA 300
Db 625 AAATTATCAGAGATGAGAAACTGAAGTACAGGACGACATGAATCTCTTAATTCAGAA 684
QY 301 TGGGATGCTCAGGGTACGACGAGAAACCAAGCAATTTACATAGATTTTAATG 360
Db 685 TGGGATGCTCAGGGTACGACGAGAAACCAAGCAATTTACATAGATTTTAATG 744
QY 361 GATCTCCAGAAATC-GAAATGGAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 419

```

```

Db 745 GATCTCCAGATCAGAAACCTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 804
QY 420 ACAAGGAAATGGAGGAAGAGCCCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 479
Db 805 ACAAGGAAATGGAGGAAGAGCCCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 864
QY 480 CAACAAATAGAGTCTTCAAGAGATCTAGAAAGATCTAGAAAGATCTAGAAAGATCTTCTC 539
Db 865 CAACAAATAGAGTCTTCAAGAGATCTAGAAAGATCTAGAAAGATCTAGAAAGATCTTCTC 924
QY 540 ACTCACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 599
Db 925 ACTCACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 984
QY 600 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGC 659
Db 985 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGC 1044
QY 560 TGGGTTCTTTTACAGACATCTCTCAAAATGGCAAGCTCTTACTGAGAACAGTGGCTT 719
Db 1045 TGGGTTCTTTTACAGACATCTCTCAAAATGGCAAGCTCTTACTGAGAACAGTGGCTT 1098
QY 720 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCGAGTGAAACAGATTTCACACAACTGGCTTT 779
Db 1099 ----- 1098
QY 780 AAGATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTTA 839
Db 1099 -----GTTTTAAAGCGGATCTTA 1116
QY 840 GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTACTCAAAACAGATCTTCTTTCAACA 899
Db 1117 GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTACTCAAAACAGATCTTCTTTCAACA 1176
QY 900 CTGAAGAAATAGTCAAGTACCCAGAGACCGAAGCATGGCTGGATTAACCTTTGCCGGTGT 959
Db 1177 CTGAAGAAATAGTCAAGTACCCAGAGACCGAAGCATGGCTGGATTAACCTTTGCCGGTGT 1236
QY 960 TGGGATTAATTTAGTCCAAAACCTTGAAGAGTACAGCACAGA 1002
Db 1237 TGGGATTAATTTAGTCCAAAACCTTGAAGAGTACAGCACAGA 1279
RESULT_4
CB228986 728 bp mRNA linear EST 10-FEB-2003
LOCUS AGENCOURT_1149247 NICHHD_Rh_Ov1 Macaca mulatta cDNA clone
DEFINITION IMAGE:6884820 5', mRNA sequence.
ACCESSION CB228986
VERSION CB228986.1 GI:28280564
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
Cercopithecinae; Macaca.
REFERENCE
1 (bases 1 to 728)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-i@mail.nih.gov
Tissue Procurement: Dr. Elliot Spindel
cDNA Library Preparation: CLONTECH
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC3135 row: e column: 11
High quality sequence stop: 583.

```

```
FEATURES
  source
    Location/Qualifiers
      1..728
        /organism="Macaca mulatta"
        /mol_type="rRNA"
        /db_xref="taxon:9544"
        /clone="IMAGE:6884820"
        /tissue_type="Ovary"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="NICHDRh_Ov1"
        /notes="Organ: ovary; Vector: pDNR-LIB; Site 1: Sfi I;
        Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dT.
        Average insert size 1.0-4.0 kb. Tissue pooled from
        pre-pubertal, post pubertal sn menopausal monkeys.
        Constructed by Clontech. Note: this is a NICHDR Library."
      ORIGIN
        Query Match      29.2%; Score 585.2; DB 14; Length 728;
        Best Local Similarity 96.9%; Pred. No. 3.2e-127;
        Matches 618; Conservative 0; Mismatches 18; Indels 2; Gaps 2;

QY 195 GGTGTAATATTCTCAATTTGGAAAGTAGCTGATTGGAACAGGAAATTTATCAGAAGA 254
Db 1 GGTGTAATATTCTCAATTTGGAAAGTAGCTGATTGGAACAGGAAATTTATCAGAAGA 60

QY 255 TGAAGAACTGAGTACAGACAGAGATGATCTCCTAATTCAGATGGGAATGCTCAG 314
Db 61 TGAAGAACTGAGTACAGACAGAGATGATCTCCTAATTCAGATGGGAATGCTCAG 120

QY 315 GGTAGCTAGCATGCGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATC- 373
Db 121 GGTAGCTAGCATGCGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCA 180

QY 374 GAAACTGAAAGATTGAATGACTGGCTTACAAACACAGAAAGAAAGAACAGGAAATGGA 433
Db 181 GAAACTGAAAGATTGAATGACTGGCTTACAAACACAGAAAGAAAGAACAGGAAATGGA 240

QY 434 GGAAGACCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACATAAGGT 493
Db 241 GGAAGACCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACATAAGGT 300

QY 494 GTTCAAGAAAGATCTAGAACAAAGCAAGTCAGGTCATTTCTCACTCACATGGTGT 553
Db 301 GTTCAAGAAAGATCTAGAACAAAGCAAGTCAGGTCATTTCTCACTCACATGGTGT 360

QY 554 GGTAGTTGATGATCTAGTGGAGATCAGCAACTGCTGTTTGGAGAACCACTTAAAGT 613
Db 361 GGTAGTTGATGATCTAGTGGAGATCAGCAACTGCTGTTTGGAGAACCACTTAAAGT 420

QY 614 ATTGGAGATCGATGGGCAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACA 673
Db 421 ATTGGAGATCGATGGGCAACATCTGTAGATGGACAGAAAGACCGCTGGGTTCTTTTACA 480

QY 674 AGACATCTTCTCAATGGCAAGCTCTTACTGAAGAACAGTGCCTTTTATGTCATGGCT 733
Db 481 AGACATCTTCTCAATGGCAAGCTCTTACTGAAGAACAGTGCCTTTTATGTCATGGCT 540

QY 734 TTCAGAAAAAGAGATGACGTGAACCAAGATTACACAACT-GGCTTTAAAGATCAAAATG 792
Db 541 TTCAGAAAAAGAGATGCTGTGAACCAAGATTACACAACTGGCTTTTAAAGATCAAAATG 600

QY 793 AATGTTATCAAGTCTTCAAAAACCTGGCCGTTTAAAA 830
Db 601 AATGTTATCAAGTCTTCAAAAACCTGGCCGTTATGAA 638

RESULT 5
AY399455
LOCUS
DEFINITION
  Mus musculus HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
  genomic survey sequence.
ACCESSION
  AY399455
VERSION
  AY399455.1 GI:39755444
KEYWORDS
  GSS.
```

```
Db 931 ACTCATGTTAGTGGTTGATGAATCCAGCGGTGATCATGCCACAGCTCTTTGGAA 990
Qy 600 GAACAACCTTAAGTATTTGGGAGATCGATGGGCAAAATCTGTAGATGGACAGACCGC 659
Db 991 GAACAACCTTAAGTATTTGGGAGATCGATGGGCAAAATCTGTAGATGGACAGACCGC 1050
Qy 660 TGGGTTCTTTTACAAGACATCTTCTCAAAATGCGAAGCTTCTTACTGAAGAACAGTGCCTT 719
Db 1051 TGGATTTCTTTTACAAGATATCTTCTTAAATGCGCAGCATTTTACTGAAGAACAG----- 1104
Qy 720 TTTAGTGCATGGCTTTTCAAGAAAAGAGATGCAAGTGAACAAGATTACACAACTGGCTTT 779
Db 1105 ----- 1104
Qy 780 AAGATCAAAATGAATGTTATCAAGTCTTCAAAACTGGCGTTTAAAGCGGATCTA 839
Db 1105 -----ACTTTAAATAATAGATCTA 1122
Qy 840 GAAAGAAAAGCAATCCATGGGCAACTGTATTCATCACTCAAAACAGATCTTCTTCAACA 899
Db 1123 GAAAGAAAAGCAATCCATGGGCAACTGTATTCATCACTCAAAACAGATCTTCTTCAACA 1182
Qy 900 CTGAAGAATAAGTCACTGAGCCCAAGAGACGGAAGCATGGCTGGATAACTTTCCCGGTGT 959
Db 1183 CTGAAGAATAAGTCACTGAGCCCAAGAGACGGAAGCATGGCTGGATAACTTTCCCGGTGT 1242
Qy 960 TGGGTAATTTAGTCCAAAACCTTGAAGAGAGTACAGCAGACA 1002
Db 1243 TGGGACAATTTAAACCCAAAACCTTGAAGAGAGTTCAGCAGACA 1285
```

RESULT 6

```
CD653550
LOCUS CD653550 777 bp mRNA linear EST 18-JUN-2003
DEFINITION AGENCOURT_14553000 NIA Human H1 Embryonic Stem Cell cDNA Library
(Long) Homo sapiens cDNA clone IMAGE:30426742 5', mRNA sequence.
ACCESSION CD653550
VERSION CD653550.1 GI:31892067
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 777)
AUTHORS NTH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaabs-x@mail.nih.gov
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
CDNA Library Preparation: Yulan Piao and Minoru Ko
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC c lone distribution information
can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: NDAM513 row: a column: 23
High quality sequence stop: 715.
Location/Qualifiers
1. 777
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30426742"
/tissue_type="Embryonic stem cells"
/cell_line="WA01"
/lab_host="DHI0B (71 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library
(Long)"
```

FEATURES

source

/note="Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEV feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTF, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and BOMES. When confluent (18-10 days after plating), the ES cells from 4 X 6 cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with TRIzol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGAGTCTTCTAGATCGGAGCGGCCCTTTTCTTTT-3'] from 3-4% of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lOne-linker L1-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The average insert size is about 3.6kb."

ORIGIN

```
Query Match 28.4%; Score 569; DB 14; Length 777;
Best Local Similarity 100.0%; Pred. No. 2.2e-123; Mismatches 0; Indels 0; Gaps 0;
Matches 569; Conservative 0;

Qy 1433 CCAGGAGCTCAAAACACTTGTCTGGGACCATCCAAAATGACAGAGCTTACAGTCTTT 1492
Db 137 CCAGGAGCTCAAAACACTTGTCTGGGACCATCCAAAATGACAGAGCTTACAGTCTTT 196
Qy 1493 AGCTGACCTGAATAATGTCAGATCTCAGCTTATAGACTGCCATCAAACTCCGAGACT 1552
Db 197 AGCTGACCTGAATAATGTCAGATCTCAGCTTATAGACTGCCATCAAACTCCGAGACT 256
Qy 1553 GCAGAGGCCCTTTGCTTGGATCTCTTTGAGCCCTGTGAGTGCATGTGATGCTTGGACCA 1612
Db 257 GCAGAGGCCCTTTGCTTGGATCTCTTTGAGCCCTGTGAGTGCATGTGATGCTTGGACCA 316
Qy 1613 GCACACCTCAAGCAAAATGACAGGCCCATGGATATCTTCGAGATTTATTAATTTGTTGAC 1672
Db 317 GCACACCTCAAGCAAAATGACAGGCCCATGGATATCTTCGAGATTTATTAATTTGTTGAC 376
Qy 1673 CACTATTATGACCGCTCGAGCAAGACACAAATTTGGTCAACGCTCCTCTCTGCGT 1732
Db 377 CACTATTATGACCGCTCGAGCAAGACACAAATTTGGTCAACGCTCCTCTCTGCGT 436
Qy 1733 GGATATGTCTGAACTGGCTGTGATGATTTATGATACGGGACGACAGGAGGATCCG 1792
Db 437 GGATATGTCTGAACTGGCTGTGATGATTTATGATACGGGACGACAGGAGGATCCG 496
Qy 1793 TGCCTCTCTTTTAAACTGGCATCTTTCCCTGTGTAAAGACATTTTGAACACAAAGTA 1852
Db 497 TGCCTCTCTTTTAAACTGGCATCTTTCCCTGTGTAAAGACATTTTGAACACAAAGTA 556
Qy 1853 CAGATACCTTTTCAAGCAAGTGCAGATTTCAACAGATTTTGAACAGCGAGGCTGGG 1912
Db 557 CAGATACCTTTTCAAGCAAGTGCAGATTTCAACAGATTTTGAACAGCGAGGCTGGG 616
Qy 1913 CTTCTCTCTGATGATTTCTCAAAATTCAGACAGTGGGTGAAGTTGCATCTCTTTGG 1972
Db 617 CTTCTCTCTGATGATTTCTCAAAATTCAGACAGTGGGTGAAGTTGCATCTCTTTGG 676
```

```
QY 1973 GGCAGTAACATTGAGCCAAAGTTCGGGA 2001
|
|
|
Db 677 GGCAGTAACATTGAGCCAAAGTTCGGGA 705

RESULT 7
AL556247 824 bp mRNA linear EST 31-MAY-2003
DEFINITION AL556247 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CSODK001YB17 5-PRIME, mRNA sequence.
ACCESSION AL556247
VERSION AL556247.2 GI:31278051
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 824)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT On Feb 15, 2001 this sequence version replaced gi:12898746.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitrogenCorporation 1600
Faraday Avenue Genoscope sequence ID : CSODK001CA09Qp1.
FEATURES
Location/Qualifiers
1..824
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODK001YB17"
/cell_type="HELA"
/clone_lib="HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 28.4%; Score 569; DB 9; Length 824;
Best Local Similarity 100.0%; Pred. No. 2.2e-123;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1433 CCAGGAGCTCAACAACTTGTGGGACCATCCAAAATGACAGAGCTCTACCAAGTCTTT 1492
|
|
|
Db 195 CCAGGAGCTCAACAACTTGTGGGACCATCCAAAATGACAGAGCTCTACCAAGTCTTT 254

QY 1493 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCAATCCGAAGACT 1552
|
|
|
Db 255 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCAATCCGAAGACT 314

QY 1553 GCAGAGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGCTGCATGTGATCCCTTGACCA 1612
|
|
|
Db 315 GCAGAGGCCCTTTGCTTGGATCTCTGAGCCTGTGAGCTGCATGTGATCCCTTGACCA 374

QY 1613 GCACAACTCAAGCAAAATGACAGCCCATGGATATCCTCAGATTATTAATTTGTGAC 1672
|
|
|
Db 375 GCACAACTCAAGCAAAATGACAGCCCATGGATATCCTCAGATTATTAATTTGTGAC 434

QY 1673 CACTATTATGACGCCCTGAGCAAGAGCAACAATTTGGTCAAGTCCTCTCGGT 1732
|
|
|
Db 435 CACTATTATGACGCCCTGAGCAAGAGCAACAATTTGGTCAAGTCCTCTCGGT 494

QY 1733 GGATATGTCGTAACCTGGCTGTGATGTTTATGATACGGGACGACGAGGAGGATCCG 1792
|
|
|
Db 495 GGATATGTCGTAACCTGGCTGTGATGTTTATGATACGGGACGACGAGGAGGATCCG 554
```

```
QY 1793 TGCTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAAGTA 1852
|
|
|
Db 555 TGCTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGACAAGTA 614

QY 1853 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTACCAAGCGAGCTGGG 1912
|
|
|
Db 615 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTACCAAGCGAGCTGGG 674

QY 1913 CCTCCTTCTGATGATTTCTATCCAAATCCCAAGACAGTTGGTGAAGTTGCATCTCTTGG 1972
|
|
|
Db 675 CCTCCTTCTGATGATTTCTATCCAAATCCCAAGACAGTTGGTGAAGTTGCATCTCTTGG 734

QY 1973 GGGCAGTAACATTGAGCCAAAGTTCGGGA 2001
|
|
|
Db 735 GGGCAGTAACATTGAGCCAAAGTTCGGGA 763

RESULT 8
BX365572 1098 bp mRNA linear EST 05-MAY-2003
DEFINITION BX365572 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CSODK001YB21 5-PRIME, mRNA sequence.
ACCESSION BX365572
VERSION BX365572.1 GI:30366927
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1098)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitrogenCorporation 1600
Faraday Avenue Genoscope sequence ID : CSIAK001ZB1Qp1.
FEATURES
Location/Qualifiers
1..1098
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODK001YB21"
/cell_type="HELA"
/clone_lib="HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 28.4%; Score 567.4; DB 13; Length 1098;
Best Local Similarity 99.8%; Pred. No. 5.6e-123;
Matches 568; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1433 CCACGAGACTCAAAACAACTTGTGGGACCATCCAAAATGACAGAGCTCTACCAAGTCTTT 1492
|
|
|
Db 204 CCACGAGACTCAAAACAACTTGTGGGACCATCCAAAATGACAGAGCTCTACCAAGTCTTT 263

QY 1493 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCAATCCGAAGACT 1552
|
|
|
Db 264 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCAATCCGAAGACT 323

QY 1553 GCAGAGGCCCTTTGCTTGGATCTCTTGGATCTCTGAGCCTGTGAGCTGCATGTGATCCCTTGACCA 1612
|
|
|
Db 324 GCAGAGGCCCTTTGCTTGGATCTCTTGGATCTCTGAGCCTGTGAGCTGCATGTGATCCCTTGACCA 383
```


Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 23 row: 9 column: 04
Seq primer: M13Rpl reverse primer (AB1).

FEATURES

Location/Qualifiers
1..620
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="he23g04"
/tissue_type="Retina"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Retina cDNA (Un-normalized,
unamplified): hc/he"
/notes="Organ: Eye; Vector: pSPORT1; Neural retina tissue
was dissected from two 80 year old donors with no observed
eye disease. 100ug of total RNA was used for library
construction. A directionally cloned cDNA library in the
pSPORT1 vector (Life Technologies) was constructed at
Bioserve Biotechnology (Laurel MD) essentially following
the protocols of the SuperScript Plasmid System full
details of which are contained in the manufacturer's
instruction manual (<http://www.lifetech.com/>). First
strand synthesis was carried out using a Not I
primer-adaptor
[5'-pACAGTCTAGATCGGAGCGCGCC(T)15-3']. EST analysis
was performed on the unamplified library at the NIH
Intramural Sequencing Center (NISC)."

ORIGIN

Query Match 27.9%; Score 558; DB 13; Length 620;
Best Local Similarity 100.0%; Pred No. 8,1e-12; Indels 0; Gaps 0;
Matches 558; Conservative 0; Mismatches 0;
QY 1444 AACAACTTGCTGGGACCAATCCAAATGACAGAGCTTACAGTCTTTAGTGACCTGA 1503
DB 1 AAACAACTTGCTGGGACCAATCCAAATGACAGAGCTTACAGTCTTTAGTGACCTGA 60
QY 1504 ATATGTCGATCTCAGCTTATAGGACTGCCATGAATCCGAACTCCGAGACTGCAGAGGCC 1563
DB 61 ATATGTCGATCTCAGCTTATAGGACTGCCATGAATCCGAACTCCGAGACTGCAGAGGCC 120
QY 1564 TTTGCTTGATCTCTGAGCTCTGAGCTGCATGTGATGCTTGGACCAACCACTCA 1623
DB 121 TTTGCTTGATCTCTGAGCTCTGAGCTGCATGTGATGCTTGGACCAACCACTCA 180
QY 1624 AGCAAAATGACAGCCCATGATATCTGCGAGATTATTAATGTTTGACCACTATTATG 1683
DB 181 AGCAAAATGACAGCCCATGATATCTGCGAGATTATTAATGTTTGACCACTATTATG 240
QY 1684 ACCGCTGGAGCAGACGACCAAAATTTGCTCAACGTCCTCTCTCGTGGATATGTC 1743
DB 241 ACCGCTGGAGCAGACGACCAAAATTTGCTCAACGTCCTCTCTCGTGGATATGTC 300
QY 1744 TGAACCTGGCTGCTGAATTTATGATACGGGACGAA CAGGAGGATCCGTGTCTGTCT 1803
DB 301 TGAACCTGGCTGCTGAATTTATGATACGGGACGAA CAGGAGGATCCGTGTCTGTCT 360
QY 1804 TTAACCTGGCATCTTCCCTGTGAAGCACAATTTGGAAGCAAGTACAGTACCTTT 1863
DB 361 TTAACCTGGCATCTTCCCTGTGAAGCACAATTTGGAAGCAAGTACAGTACCTTT 420
QY 1864 TCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCTCTCTGC 1923
DB 421 TCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCGCAGGCTGGGCTCTCTGC 480
QY 1924 ATGATCTATCCAAATCCAAAGCAGTTGGGTGAAGTTCATCTCTTTGGGGCAGTAACA 1983

Db 481 ATGATTCTATCCAAATCCAAAGACAGTTGGGTGAAGTTGCATCTTTGGGGCAGTAACA 540
QY 1984 TTGAGCCCAAGTGTCGGA 2001
DB 541 TTGAGCCCAAGTGTCGGA 558

RESULT 11

BG719710
LOCUS BG719710 770 bp mRNA linear EST 08-MAY-2001
DEFINITION 602690430F1 NIH_MGC_97 Homo sapiens CDNA clone IMAGE:4822807 5',
mRNA sequence.

ACCESSION BG719710
VERSION BG719710.1 GI:13998897

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (Bases 1 to 770)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

CDS Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: LLAM10731 row: 1 column: 08

High quality sequence stop: 767.

Location/Qualifiers

1..770

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4822807"

/lab_host="DH10B"

/clone_lib="NIH_MGC 97"

/note="Organ: testis; Vector: pBluescriptR (modified

pBluescript KS+), Site_1: BamHI; Site_2: SalI-XhoI

(GTCGAG); Cligo-dr primed using primer

5'-TTTTTTTTTTTTTTN-3', size-selected for average

insert size 2.2 Kb and normalized to ROT 5. This is a

primary library enriched for full-length clones and

constructed using the Cap-trapper method (Carninci, in

preparation) library constructed by M. Brownstein

(NIH/NHGRI, National Institutes of Health). Note: this is

a NIH_MGC Library."

ORIGIN

Query Match 26.6%; Score 532.2; DB 12; Length 770;
Best Local Similarity 98.1%; Pred. No. 1.1e-114;
Matches 560; Conservative 0; Mismatches 8; Indels 3; Gaps 2;
QY 1433 CCACGAGACTCAACACTTCTGCGACCATCCCAAAATGACAGGCTTACCAAGTCTTT 1492
DB 150 CCACGAGACTCAACACTTCTGCGACCATCCCAAAATGACAGGCTTACCAAGTCTTT 209
QY 1493 AGCTGACCTGAATATGTGACATTCAGATTCTAGGACTGCCATGAACTCCCAAGACT 1552
DB 210 AGCTGACCTGAATATGTGACATTCAGATTCTAGGACTGCCATGAACTCCCAAGACT 269
QY 1553 GCAGAAGGCCCTTTGCTTGGATCTCTTGGAGCTCTCAGCTGTCATGTGATGCTTGGACCA 1612
DB 270 GCAGAAGGCCCTTTGCTTGGATCTCTTGGAGCTCTCAGCTGTCATGTGATGCTTGGACCA 329
QY 1613 GCACAA--CCTCAGCAAAATGACAGCCCATGATATCTTGCAGATTATTAATGTTG 1670

Db 330 GCACATCTCAAGTCAGATGACCCAGCCCATGGATATCTCGAGATTATTAAATGTTTG 389

Qy 1671 ACCACTATTTATGACCGCTGGAGGACAGACAAATTTGGTCAACCTCCTCTCTGC 1730

Db 390 ACCACTATTTATGACCGCTGGAGGACAGACAAATTTGGTCAACCTCCTCTCTGC 449

Qy 1731 GTGATATGCTGTGAACCTGCTGCAATGTTTATGATACGGACGACGAGGAGATC 1790

Db 450 GTGATATGCTGTGAACCTGCTGCAATGTTTATGATACGGACGACGAGGAGATC 509

Qy 1791 CGTGTCTCTCTTTAAACTGGCATCATTTCCCTGTGTAAGACACATTTGGAAGACAA 1850

Db 510 CGTGTCTCTCTTTAAACTGGCATCATTTCCCTGTGTAAGACACATTTGGAAGACAA 569

Qy 1851 TAGCATACCTTTTCAGCAAGTGGCAAGTTCACAGGATTTTGTGACGAGCGAGGCTG 1910

Db 570 TAGCATACCTTTTCAGCAAGTGGCAAGTTCACAGGATTTTGTGACGAGCGAGGCTG 628

Qy 1911 GGCTCTCTCTGTCATGATTTCTATCCAAATTCAGACAGTGGTGAAGTTCATCTCTTT 1970

Db 629 GGCTCTCTCTGTCATGATTTCTATCCAAATTCAGACAGTGGTGAAGTTCATCTCTTT 688

Qy 1971 GGGGGCAGTAACTAGGCAAGTGGTCCCGA 2001

Db 689 GGGGGCAGTAACTAGGCAAGTGGTCCCGA 719

RESULT 12

AK087829

LOCUS

DEFINITION

AK087829 Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:E330026B12 product:dystrophin, muscular dystrophy, full insert sequence.

ACCESSION

AK087829

VERSION

AK087829.1 GI:26352777

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 9279253

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itch, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 20499374

3 Shibata, K., Itch, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kutsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium Nature 409, 685-690 (2001) 11076861

5 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium Nature 409, 685-690 (2001) 11076861

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

FEATURES

Location/Qualifiers

1..1298

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:E330026B12"

/db_xref="MGI:2426619"

/db_xref="taxon:10090"

/clone="E330026B12"

/sex="female"

/tissue="ovary"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="2 days pregnant adult"

120..1142

/notes="unnamed protein product: dystrophin, muscular dystrophy (MGI:94909, GI:NM_007868, evidence: BLASTN, 100%, match=999)

putative"

/codon_start=1

/protein_id="BAC40019.1"

/db_xref="GI:26352778"

/translation="MREHLKGHTOTTCWDHPKWTLYOSLADLNVRFSAYATAMKL RRLQKALCDLLSLSAACDALDQNLKQNDQNDLQIINCLTITVDLQEQHNVLN VPLVCMDLNLNLVDYDGTGTRIRLVSKTIGLISCKAHLEKRYLFKQVASSIGF CDQRGLGLLHDSIQIPQGEVAVSGGNSIEPSCFQFANNNKPEISAAFLDWMR LQFQSMVLPVLRVAATCAQKCIKECIPIGRYSRLKHFNYDICOCSFFSG RVAGKHGMVPMVEYCTPTTSGEDVRDFAKVLKNKFKRYFAKHPRMGYLPQVTLE GDNNET"

CDS

Query Match 25.6%; Score 511.4; DB 11; Length 1298;

Best Local Similarity 93.7%; Pred. No. 9.7e-110;

Matches 533; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 1433 CCAGAGACTCAAAACACTTGTCTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTT 1492

Db 140 CCAGAGACCAAAACACTTGTCTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTT 199

Qy 1493 AGCTGACCTGAATAATGTCAGATTCTTCAGCTTAGAGCTGCCATGAACTCCGAGACT 1552

This Page Blank (uspto)